Advertising, Imperfect Information

and the Effect of

Learning on Consumer Behaviour

by

Ian Tonks

Thesis submitted to the University of Warwick for the degree of Doctor of Philosophy, July 1983.

The research in this thesis was carried out in the Department of Economics, University of Warwick; School of Economics and Social Studies, University of East Anglia; and Department of Economics, University of Exeter.

Table of Contents

Acknowledgements

This thesis was started in October 1979. I was financed for two years full time research by a Social Science Research Council Linked Studentship at the University of Warwick. The work continued part time while I was employed by the University of East Anglia from October 1981 to September 1982, and has been completed since being employed as the Esmee Fairburn Lecturer in Finance at the University of Exeter from October 1982.

A version of Chapter 4 was given as a paper at the Association of University Teachers of Economics Conference at Surrey in April 1982, and appeared in the conference edition of the Economic Journal. Part of Chapter 4 has also been accepted for publication in the Review of Economic Studies.

I owe a debt of gratitude to an enormous number of people who have commented at various stages, read through parts and made valuable points: Norman Ireland, my supervisor at Warwick, Paul Weller, Avinash Dixit, Clive Fraser, Keith Cowling, Steve Davies, Mike Morris, Jane Black, John Matatko, George Bulkley, Frank Oliver, Desmond Corner, Andy Snell, Ben Lockwood, Geoff Stewart; members of the Data Processing Unit at Exeter for providing computer programming assistance, especially Ray Burnley who wrote the stochastic dominance programme in Chapter 8. I would also thank all colleagues in the economics departments at Warwick, East Anglia and Exeter for providing an environment in which it was possible to carry out this work.

My final thanks go to Vi Palfrey and Melanie Hall who typed up the thesis.

Glossary of Notation

A	matrix of unknown parameters in chapter 3
a,b	fixed coefficients in household technology in chapter 4
a,b	parameters in inverse pareto distribution in chapter 6
a,b	parameters of time profile for individual in chapter 7
c	characteristics
d,e	adjustment costs in chapter 4
b_o	observed sample statistic in chapter 6
i	social rate of discount and market rate of interest
k	number of advertising messages received in chapter 3
L	loss function in chapter 8
M	income
m	random variable with standard normal distribution
P	prices
r	known degree of precision of log normal random variable
s	quantity of good complementary to x and y in chapter 4
τ,t	time periods
T	final time period in environment
EU	discounted sum of expected utility
u	utility per period
V	indirect utility function
$\underset{\sim}{X}$	vector of observations
x	quantity of new good
y	quantity of old good
z	unknown parameter
α	unknown parameter in chapter 3
α,β	parameters of utility function in chapter 5
γ	coefficient of adjustment in adaptive learning process
ε	random variable, representing product quality

$\bar{\varepsilon}$ observed level of product quality

μ parameter supplied by the firm

μ_o mean of subjective beliefs

ϕ consumer's degree of confidence

π implicit prices in chapter 2

π discrete probabilities in chapter 3

ρ deception element of an advertising message in chapter 3

η difference between expected and actual product quality in chapter 7

χ elasticity of demand in chapter 7

Summary

This thesis incorporates a role for advertising in a neoclassical model of consumer behaviour in an environment with imperfect information. Advertising is seen as providing an initial set of parameters in the consumer's subjective beliefs. The thesis considers the role of learning through experience over time, as an alternative source of information. The ability to learn is shown to affect consumer behaviour under a number of assumptions about the state of the environment. The models under consideration are complex, and analytical solutions are difficult to disentangle. For this reason the thesis makes use of numerical analysis to provide solutions for the specific parameter values assumed. The desire of a consumer to gain information is shown to generate a pattern of purchases of a new good which accords with the empirical evidence. Over time, we find that the initial advertising statement declines in importance as a component of current beliefs, as the consumer places greater reliance on his own experience. We contrast different learning mechanisms in an attempt to find the most efficient process for gaining complete information, and conclude that no one learning process is always dominant, but depends on the parameters in the environment.

Chapter 1

Introduction

This thesis is concerned with examining consumer behaviour over a number of time periods, when the consumer is able to learn over time about unknown elements in the environment. The aim of the thesis is to incorporate advertising into a neoclassical model of consumer behaviour and consider the effect of advertising statements on the demand for a product.

Standard microeconomic theory assumes that consumers have perfect knowledge of all available goods and prices, enabling them to compute easily their optimal demands and supplies. Once we remove the assumption of complete knowledge we enrich the structure of the model by introducing an additional tier of choices into the consumer's decision problem. Before he can decide which goods to buy, he must first of all acquire information about the available products, such as to their existence, functions, quality, location and price. The consumer must determine both the quantity and quality of the information required. The demand for information is a derived demand, since it allows the consumer to make an informed decision about the demand for a final product. The quality of the information will depend upon its source. Information may be obtained from a number of sources. A consumer could walk the streets, observing prices and testing individual products. He may read about goods in factual magazines or receive details from consumer watchdog organisations. He may find out about product characteristics on the basis of his own experience, from purchasing the good in the past. Other consumers may tell him about particular attributes of goods that they have sampled, enhancing the reputation of some goods. He will be subjected to advertising statements on television, on billboards and in newspapers, all purporting to give information about the respective product. The consumer will select various bits of information, some of which he pays for directly and others which are free, and use them to compose a set of beliefs about the unknown elements in his environment. In this thesis we suppose that the consumer has two sources of information: advertising messages and his own experience. Initially, by definition, his experience set is null, and his information set consists entirely of advertising messages and his views about the accuracy of these statements. Over time however the consumer acquires knowledge by sampling various products, and we are interested in investigating this process.

The thesis can be divided into two sections, in the next two chapters we outline our understanding of the advertising process from the viewpoint of the consumer. We then use this definition in the remaining chapters to analyse the effect of advertising and learning on consumer behaviour

Chapter 2 attempts to distinguish between the various meanings of advertising with a view to establishing the importance of the information content of an advertising message as the crucial aspect of the effect of advertising on consumer behaviour.

Firstly we look at Marshall's distinction between constructive and combative advertising, since this classification started the debate on the difference between information and persuasive advertising. We then consider the importance of product differentiation on providing a role for informative messages, and briefly summarise the psychological aspect of how an informative statement can be presented before a consumer. Product differentiation means that each good has individual attributes, which may be real or perceived. These attributes need to be identified by the consumer, and it is argued that advertising fulfils this identification role.

We go on to examine the case of advertising which creates new wants, and the difficulties that this makes for welfare comparisons. These problems can be overcome by redefining the welfare yardstick. Here it is shown that we can compare changes in an informative message with changes in product price, in terms of income and substitution effects. This analogy is used to solve a problem which is raised in Chapter 5.

Chapter 3 defines the meaning of advertising which will be used throughout the rest of the thesis. Advertising is defined as being an information generating mechanism which can be compared with other mechanisms, such as search. We point out where imperfect information may occur in a consumer's decision problem, and the value of advertising under these circumstances. The value of perfect information and of advertising is highlighted by two examples. We then suggest a way that advertising can be incorporated into a consumer's beliefs about his environment. The crucial characteristic about advertising is that it is information given by the producer of the good, who has an incentive to exaggerate its positive attributes. However, the ability of the producer to distort the truth depends upon the efficiency of the consumer in accumulating complete, unbiased knowledge.

In Chapters 4, 5 and 6, we move on to the analytics of consumer behaviour in a two-period world. We consider the problem facing a consumer when he has a choice between a new good with unknown characteristics and an old, safe reliable good. The consumers initial beliefs about the new good are given by an advertising message. The consumer is able to gain further unbiased information by purchasing the good and observing the results for use in the second period. We are interested in how much of the new good the consumer purchases, and how this optimal quantity changes as the parameters in the problem change, such as the initial advertising message and the consumers degree of confidence in the truthfulness of this message.

An underlying assumption in all the models considered is that the environment is stochastic. Product characteristics are regarded as random variables. Now this may seem an unreasonable assumption since it is likely that with mass production, products will be standardized, so that two units of the same good will be identical. However, the consumer is concerned with the perception of a good's qualities. And the same good consumed in the morning or in the evening may have a different effect on utility. Thus, the environment is randomised, not on the production side but due to the time, place and conditions under which the consumption of the good occurs. Although this model generalises to the case where product quality is itself a random variable.

In order to build models of consumer behaviour that contain the dimensions of both uncertainty and time, the structure of these models will be, by nature, complicated. Analytical solutions to optimal behaviour patterns and comparative static exercises will be difficult to define. One way round this problem of a trade-off between realism and complexity, is to consider numerical analysis of the solutions. This practice is followed throughout the thesis: the answers to various questions are given for particular values of the parameters in the problem. Thus we lose some generality in our results, but at least the results exist for specific conditions.

In Chapter 4, the underlying random variable has a normal distribution, and the consumers initial beliefs about the advertising message are also normal. In Chapter 5, the underlying distribution is log-normal, though the prior is still normal. In Chapter 6, we revert back to a symmetric distribution to describe the random variable, but the prior beliefs are assumed to have an inverse Pareto distribution. By comparing the results of these chapters, we will be able to say something about the effect of these distributions on consumer behaviour. Similarly, for simplicity, Chapters 4 and 6 assume that the consumer has a linear utility function. In order to compare the effect of risk aversion on the problem, in Chapter 5 the consumer is assumed to be risk averse; his preferences are represented by a constant elasticity of substitution utility function.

In Chapter 7, we extend the model in Chapter 4 to a many-period framework. This enables us to chart the demand for the new good over time, as information accumulates. Again we see how this pattern of purchases changes as the parameters in the problem change. We are especially interested in the speed of adjustment to the final equilibrium. That is, how long does it take the consumer to acquire sufficient information as to make an informed decision? The model generates an aggregate time profile of purchases similar to the standard diffusion curve identified in the literature. Further we state a testable hypothesis about the demand elasticities along the life-cycle of the product.

Chapter 8 attempts to compare the value of the learning from Bayesian sampling with other, less complex processes. The conclusions of the thesis are drawn together in Chapter 9.

Chapter 2

On the Importance of Distinguishing Between Informative and Persuasive Advertising

Constructive and Combative Advertising

Initially Alfred Marshall made a distinction between constructive advertising which "includes all measures designed to draw the attention of people to opportunities for buying or selling, of which they may be willing to avail themselves"(1) and combative advertising, which "obtrudes itself in the incessant iteration of the name of a product, coupled perhaps with the claim that it is of excellent quality".(2)

The contrast between these definitions is that constructive advertising is undertaken by the producer to inform the consumer of the existence of a product; it is the initial message which once observed means that the consumer can make an informed decision. Combative advertising are those same messages repeated continually which serve no useful purpose and involve social waste since "the lavish advertiser must deduct his expenses from the gross profits of his additional sales; while the rivals whom he ousts lose their gross profits".(3)

Marshall is regarding the wastefulness of advertising from the position of a consumer who has both seen the advertising statement and remembered it. The problem facing the producer is to ensure that a message is seen by as large a percentage of the population as possible. This will undoubtedly mean that the same message will be sent out more than once at different times of the day, in different locations and through different media. The greater the quantity of the same message released, the greater the probability that any one consumer will see it. Diminishing returns with respect to the number of messages may set in; that is the rate at which the proportion of the population that have seen the message increases, falls with the number of messages. The optimal number of advertising

Text cut off in original

messages for a profit maximising producer will be given by equating the cost of an additional message with the expected increase in returns from an additional proportion of the population seeing the advertising statement. A consequence of this policy is that it may well be the case that some of the population see the same message more than once.

Similarly, it may be necessary for a producer to reiterate an advertising statement because consumers' memories are not perfect, and information acquired in the past depreciates at a rate of forgetfulness which varies between consumers.

Marshall illustrates his distinction between constructive and combative advertisements: "For instance a good frontage on a leading thoroughfare; adequate space for the convenience of employees and for customers; lifts and moving staircases, etc., are all constructive so long as they do not exceed the requirements of business .... But eager rivalry often causes them to be carried to an excess, which involves social waste".

The dividing line is drawn by "the requirements of business" which is vague, but perhaps relates to the equilibrium conditions suggested earlier

For Marshall, the distinction between constructive and combative advertising is the quantity of messages. Up to some point advertising statements are constructive, but past this point continual repetition of the same statement becomes combative.

The wastefulness of combative advertising is that it simply redistributes goods between producers, a point also made by A.C. Pigou with respect to competitive advertising; "competitive advertising [is] directed to the sole purpose of transferring the demand for a given commodity from one source of supply to another".(5)

The wastefulness arises because producers are advertising the same good. Marshall and Pigou are writing against a background of competitive markets and homogeneous products. Once consumers have been made aware of the existance of an industry product, then competitive conditions dictate that the price will be the same for all firms in the industry. Advertising by firms in this situation should inform the consumers of the location of the product, and be sufficient to prevent the inception of localised monopolies.

Real and Perceived Differences in Product Quality

The development of the theory of monopolistic competition introduced the idea of an industry consisting of heterogeneous goods. Once it is recognised that different firms in the same industry produce different goods, then it can no longer be argued that advertising wastefully redistributes sales between producers. An advertising message is the means by which a producer informs the consumer about the quality and uses of his particular differentiated product. In terms of the characteristic approach to utility theory, consumers are assumed to gain utility from consuming characteristics, which are produced by goods. These characteristics come packaged in varying proportions in various goods. If individuals are not identical then we can expect that their preferences over characteristics will differ. One individual will have a preference for a characteristic which is available in one particular good, another individual will have a set of preferences which can best be satisfied by a good with a different amount of characterstics.

Product differentiation is not necessarily wasteful since it supplies consumers with a greater choice of characteristics than if there was only one homogeneous industry good. Advertising messages supply consumers with information about the quantities and proportions of the characteristics in these differentiated goods. Advertising is the means by which actual product differences can be imparted to consumers. "Product differentiation is propogated by differences in the design or physical quality of competing products, by efforts of sellers to distinguish their products through packaging, branding and the offering of auxiliary services to buyers, and by advertising the sales-promotional efforts designed to win the allegiance and custom of the potential buyer. These latter efforts may turn in part on informing the consumer of the distinctive physical properties of the individual products, and also in part on "convincing" him of their desirability, presige or superior quality".(6)

The wastefulness of product differentiation occurs because the proliferation of products means that economies of scale can not be reaped from mass production. Lancaster (1975) argues that this productive inefficiency must be offset against the gain in variety from having a multitude of goods.

Even if two goods are objectively very similar such that the loss in productive efficiency is greater than the gain from variety, then we still can not say that advertising their similar characteristics is wasteful since an Joyce (1963) notes advertising adds "psychological values to the product".(7)

Moving onto the more psychological aspects of advertising, we have so far argued that an advertising message is a piece of information sent from the producer to the consumer. The producer must ensure that the consumer comprehends the message; to achieve this the message must gain the consumer's attention and the producer must estimate how the consumer will perceive the message once read.

"Advertisements can in general only work by being seen or heard by consumers and operating on their minds, a process which in some way influences one aspect of their behaviour - the physical operation of purchasing".(8)

This more psychological approach to advertising could be regarded as persuasive, that is, the information must be presented in a persuasive way in order that it can enter the consumer's decision problem. Given that the consumer is operating in a noisy environment the information must attract his attention and be presented persuasively. According to Hicks "The attention of the consumer has to be attracted and his attention aroused. In order to perform its social function advertising has to be attractive and (let us not be afraid to say) persuasive."(9)

The Creation of New Wants and the Manifestation of Latent Wants

Up till now advertising has been regarded as providing consumers with information about unknown product characteristics, and how that information is presented. But as well as providing information about products, advertising can also be used to provide information to consumers about their own utility functions. If consumers' preferences are fixed, then advertising can be used to provide information either about parameters whose existence is known to consumer but whose value is unknown; or about parameters which are lying dormant in the consumer's utility function.

If preferences vary over time, then advertising may be responsible for these new preferences. The dividing line between advertising providing information about something that already exists and the creation of new parameters and hence new wants is very fine. If preferences do vary, then it becomes very difficult to judge whether welfare has changed for the better or worse. If advertising allows the manifestation of latent wants, then the welfare yardstick remains the same, but if advertising creates new wants then we have two standards of judgement: the new and old set of preferences.

Marshall recognised that wants may change over time: "although it is man's wants in the earliest stages of his development that give rise to his activities, yet afterwards each new step upwards is to be regarded as the development of new activities giving rise to new wants, rather than new wants giving rise to new activities".[10]

In the early stages of man's development his desire for food caused the hunter to travel. At a later stage, the invention of the airplane caused a new desire for travel by air - flying for its own sake. The introduction of a new good created a new want. The philosophical problem is to decide whether the desire for flying was a new want, or if it had also been present in stone-age man but had been lying dormant for centuries.

Similarly Pigou argues: "some advertisements serves to develop an entirely new set of wants on the part of consumers, the satisfaction of which involves a real addition to social well-being".(11) Pigou is going one stage further than Marshall, claiming that not only do advertisements manufacture new wants but that this increases consumer welfare. But this argument is not valid since we can not compare the consumer's welfare before and after the change: interpersonal comparisons or intrapersonal comparisons at different points in time are not always possible.

Marshall is making a much weaker statement: that the design of new products may stimulate new wants; though Marshall does not believe advertising can affect wants that the consumer can control: "Of course no amount of expenditure on advertising will enable anything, which the consumer can test fairly for themselves by experience (this condition excludes medicine which claim to be appropriate to subtle diseases etc.), to get a permanent hold on people unless it is fairly good relative to its price".(12) Thus a consumer's ability to learn from experience ensures that exaggerated advertising claims can be expurgated.

Although Marshall recognised the importance of advertising in creating new wants, he seemed to believe that the purpose of advertising was to "educate" consumers with respect to an undiscovered parameter in the utility function. "For the main constructive acts are often confined to discovering the existence of a latent want, and educating the public with regard to it".(13)

This educative role, was challenged by Braithwaite (1928) in a seminal article which dealt with the economic consequences of advertising. Braithwaite distinguished between 'true' selling costs, "which merely conveys information as to available supply", and advertising costs which "are incurred either by producer or middleman with a view to increasing the sale of a commodity".(14) Braithwaite argues that advertising is not educative because it does not present an unbiased set of information before the consumer: "[It] might be true ..... [that] advertisements were indeed educative ... if it put before the consumer such detailed and truthful information as would assist him in making a more correct judgement of commodities .... Advertisements are not written to help people make a reasoned choice of commodities, they are written with the object of inducing them to buy particular things, and they naturally exaggerate the uses and merits not only of the commodity but of a particular make of the commodity. Moreover, the vast majority of advertisements do not confine themselves to pointing out the uses of commodities; they make their appeal not to the reason, but to the emotions, of the consumer. Suggestion, reiteration, attractive illustration - these are all devices to induce him to buy an article without making comparisons and calculations. They certainly do not assist his judgement as to the relative satisfactions to be obtained from the different commodities or as to the relative satisfactions to be obtained from commodities and leisure .... Consequently there appears no gounds for saying that advertisement costs, which persuade the consumer to adopt this new set of judgements are conferring an educative service".(15)

Thus Braithwaite introduced the notion that advertising may "persuade" consumers to alter their preferences and provided an explicit account of how this might take effect: "Advertising expenditure as thus defined aims at increasing sales by affecting the mind of the consumer. By various appeals it induces him to change his subjective valuation of the commodity. Thus the marginal utility of the commodity is increased and the demand curve raised. Consumers are persuaded to buy more of it at the same price, the same quantity at a higher price, or even in some cases more of it at a higher price".(16)

Tintner (1952) and Basman (1956) provide a more formal treatment of how a change in a parameter in the utility function will affect the demand for a commodity.

Difficulties of Welfare Comparisons

The problem with advertising which shifts preferences is that it becomes impossible to make welfare comparisons with respect to the situation before and after the advertising message has been seen.

Neoclassical consumer theory conceives of a consumer with a preference ordering defined over all possible bundles of goods, yielding a set of indifference curves which can be assigned arbitrary values, termed utility, which must have higher values for higher indifference curves, that is utility is only defined ordinally. Consumers construct their preference ordering on the basis of the information they have before them. The consumer then maximises utility, as if it has a numerical value subject to a budget constraint and arrives at an optimum allocation between goods such that the marginal utility of a good relative to its price be equal for all goods.

Each indifference curve represents a level of utility, and hence is a measure of the consumer's well-being. Changes in the parameters in the budget constraint will move the consumer onto another, of the same set of indifference curves. The consumers standard of judgement remains fixed, by assuming a consumer prefers more of the good to less, we are able to state unambiguously whether welfare has risen or fallen. This is in contrast to the situation where the consumer changes his preference ordering. In this case the reference point used to make welfare comparisons has changed. Welfare is only defined in terms of a particular preference ordering.

Thus, if I have a preference for my house to be painted red, I can evaluate the effect on my welfare of a change in my income or a change in the price of red paint. However, if my tastes change such that I now prefer my house painted blue, can I say whether I am better off before or after the change in tastes?

Dixit and Norman (1978) attempt to answer this question, by making two welfare caloulations, they evaluate a consumers decision in terms of pre and post-advertising tastes. Thus they evaluate the utility from having my house painted blue as compared with having my house painted red in terms of my preadvertising preferences and then evaluate the same decisions in terms of my post advertising tastes. Dixit and Norman use this apparatus to look at the question of social welfare including the effects on monopoly profits of an increase in advertising.

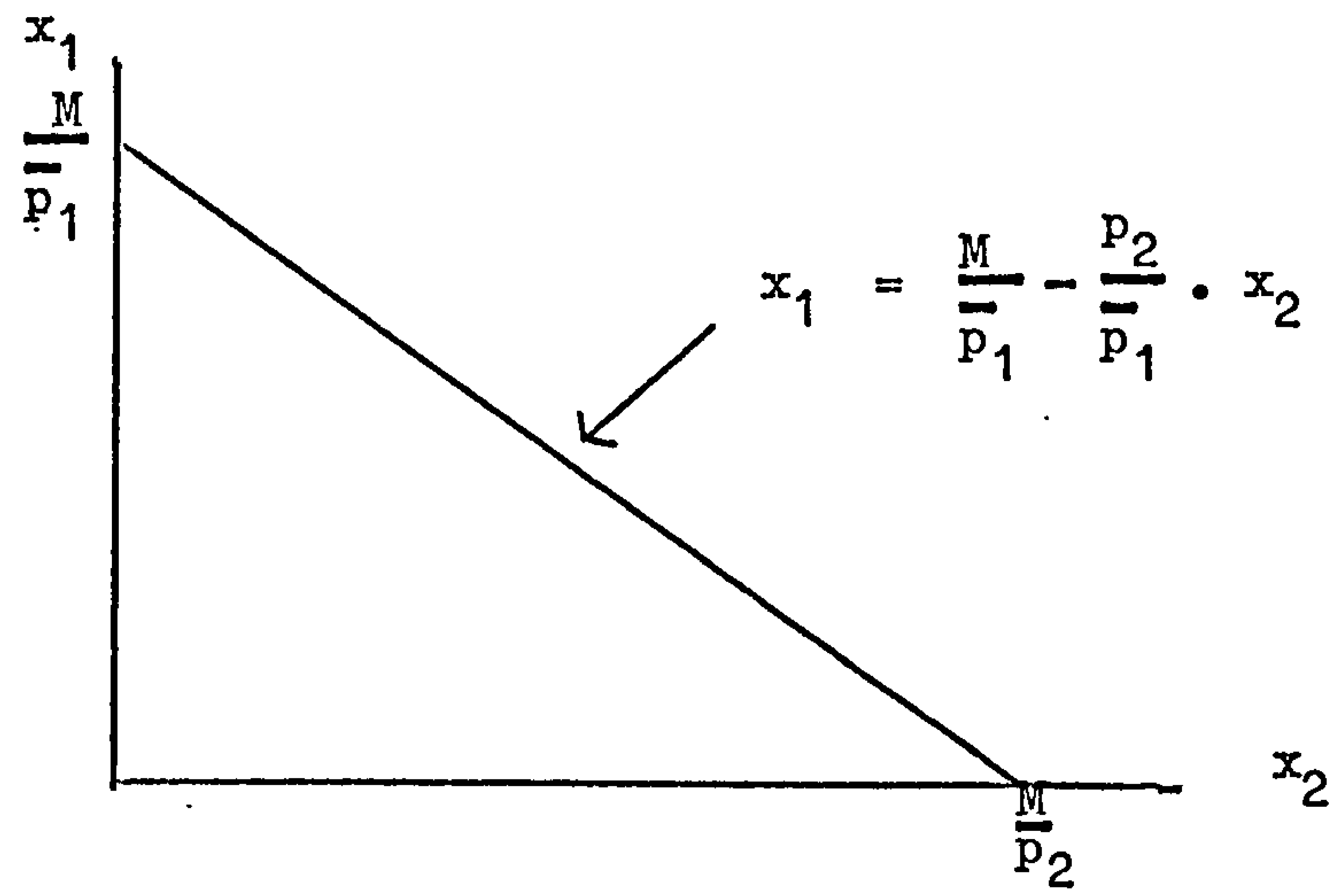
Fisher and McGowan (1979) note that the real comparison should be the utility from a red house evaluated at pre-advertising preferences and the utility from a blue house evaluated at post-advertising tastes. Dixit and Norman assume that a small amount of advertising will only have a small effect on a consumers decisions and hence only a small effect on tastes; and in effect they ignore making this comparison on the basis of its negligible value.

We can illustrate the distinction between advertising that provides information about an unknown parameter in the consumer's decision problem and advertising that shifts preferences.

Consider a consumer who has a given set of preferences between two goods x_1 and x_2, from which he is able to construct a set of indifference curves. However the consumer does not know with certainty the price of one of the goods, say x_1. He has an expectation of the price $E(p_1) = \bar{p}_1$, and constructs the budget constraint on the basis of this estimate:

$$\bar{p}_1 x_1 + p_2 x_2 = M$$

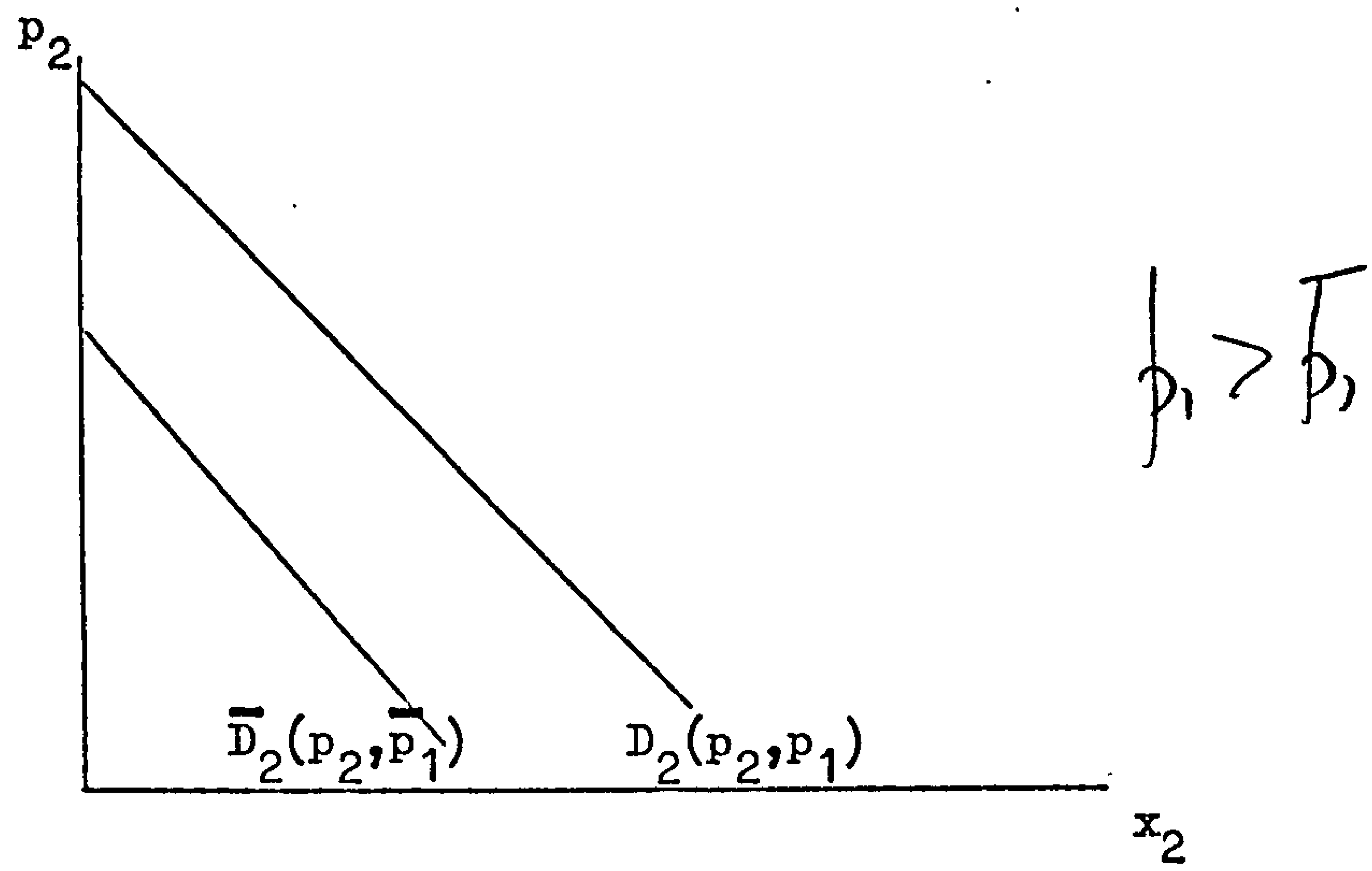
The slope of the budget line, which is essentially a subjective budget constraint, since it is what the consumer believes the budget line to be, is $\frac{p_2}{\bar{p}_1}$



Changes in relative prices will yield a series of tangency positions between the budget line and the indifference curves, from which the consumer derives his demand curve. The demand curve for the second good is derived assuming a particular price expectation for good 1.

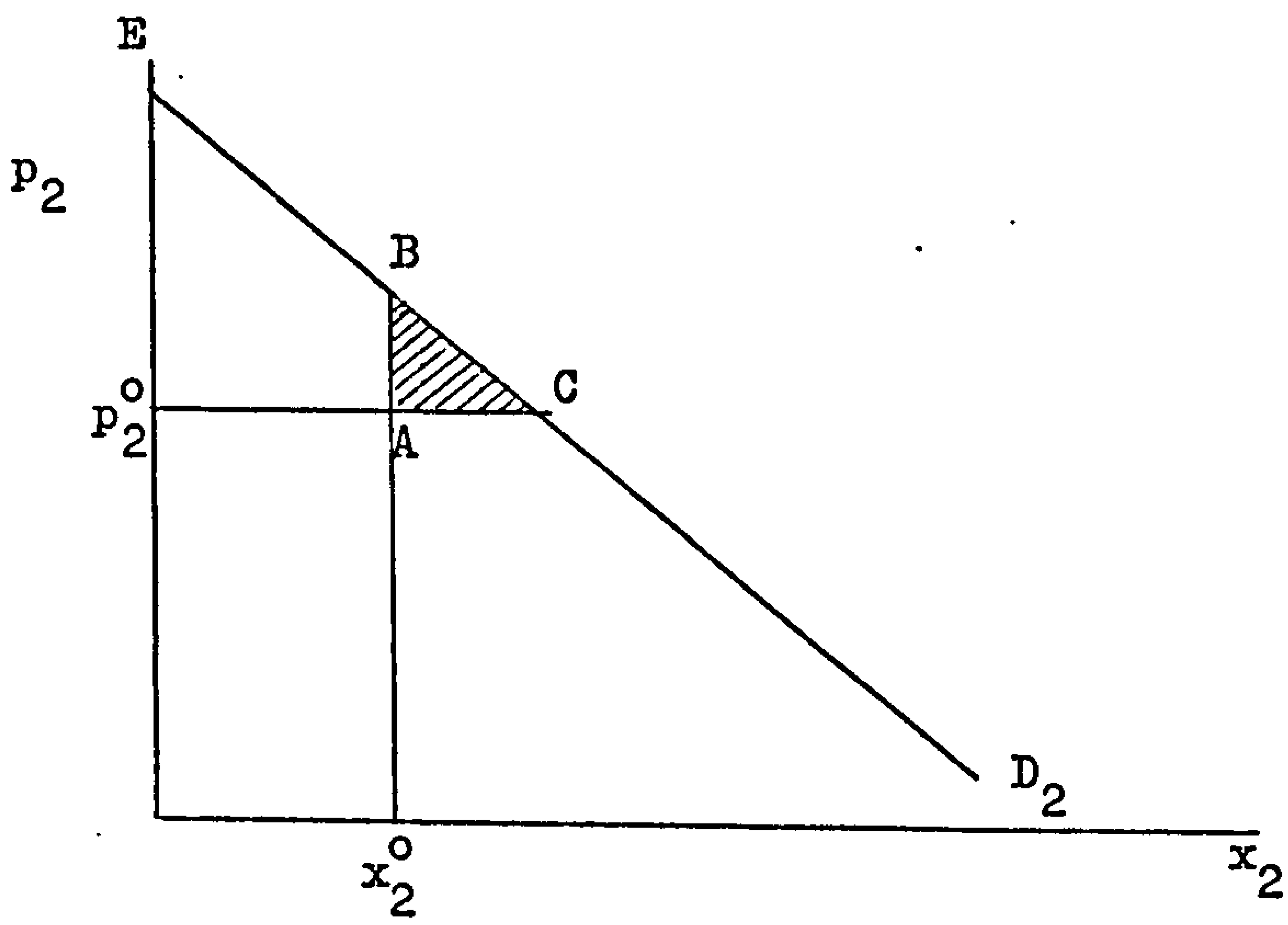
$$x_2 = \bar{D}_2 [p_2 , \bar{p}_1]$$

But suppose the actual price of good 1 is p_1 which is more expensive than the consumer believes: $p_1 > \bar{p}_1$; then if we make the reasonable assumption that $D_2(\cdot)$ is monotonically non-decreasing in the price of the first good then the true demand curve for the second good will be shifted to the right. Since if x_1 is a substitute for x_2, a rise in the price of x_1, will reduce the demand for x_1 and increase the demand for x_2 at each p_2.



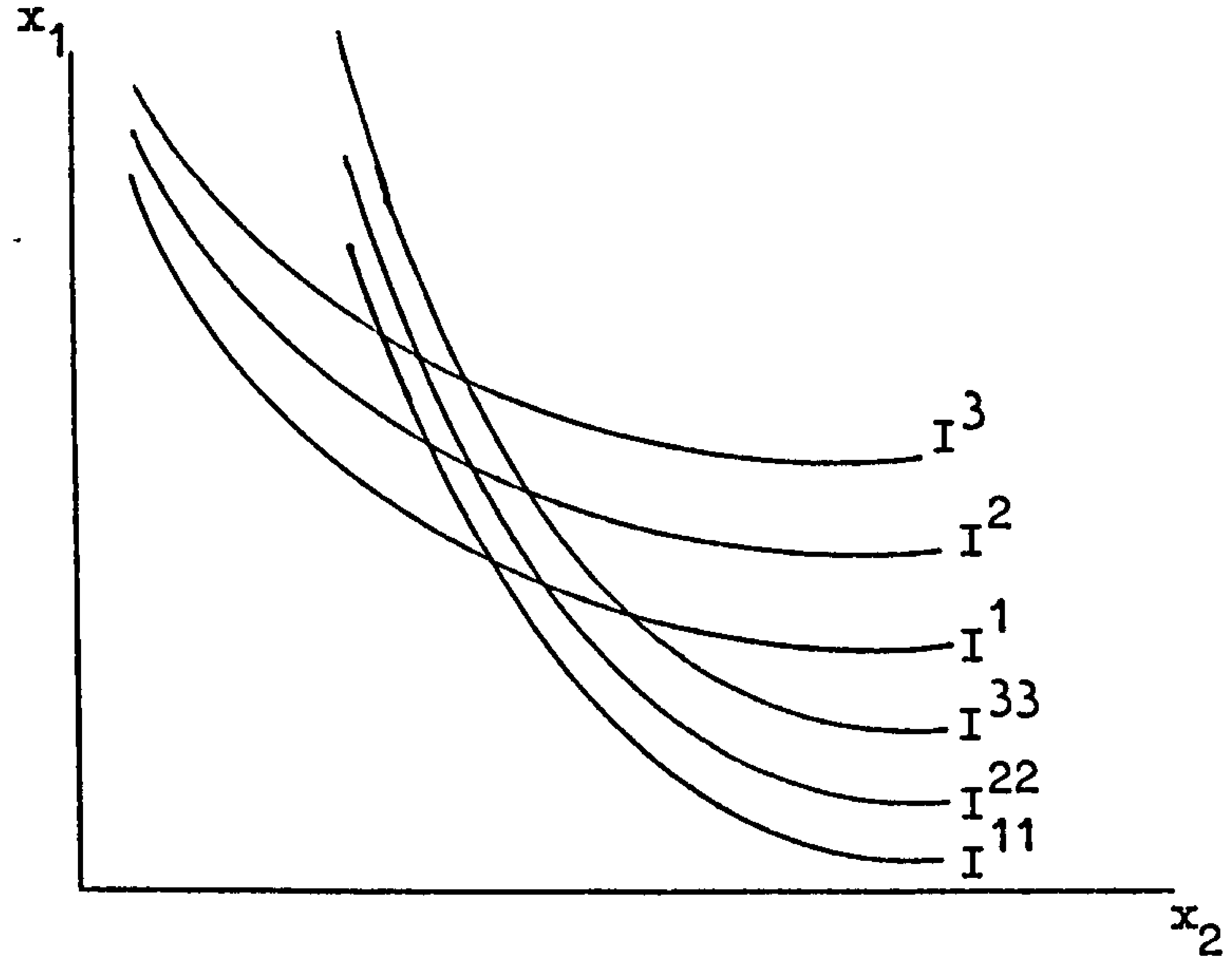
In fact for some goods, such as necessities, an increase in their price may shift the demand curve for a non-necessity such as x_2 to the left, since the income effect of the higher price outweighs the substitution effect.

D_2 (p_2 , p_1) is the true demand surve and reflects the consumer's true evaluation of the good x_2. Whereas $\bar{D}_2$ (p_2 , $\bar{p}_1$) is the perceived demand curve which the consumer bases his purchasing decision on.



At price p_2^o, the consumer demands x_2^o of the second good, since point A is a point on his perceived demand curve. His consumer surplus which is a measure of his welfare is given by the area p_2^o ABE. In purchasing x_2^o at price p_2^o, the consumer is not on his actual demand curve. The consumer can increase his welfare by increasing his consumption of x_2, until his marginal evaluation of the good is equal to its price at point C. Consumer surplus becomes p_2^o CE. The consumer can increase his welfare by the shaded area ΔABC, if he is informed of the true value of p_1. This type of advertising which informs the consumer of the value of a parameter in his decision problem unambiguously increases his welfare.

Contrast this with advertising that changes tastes. Here the advertising message affects the preferences of the consumer, and this can be represented by the set of indifference curves shifting around.



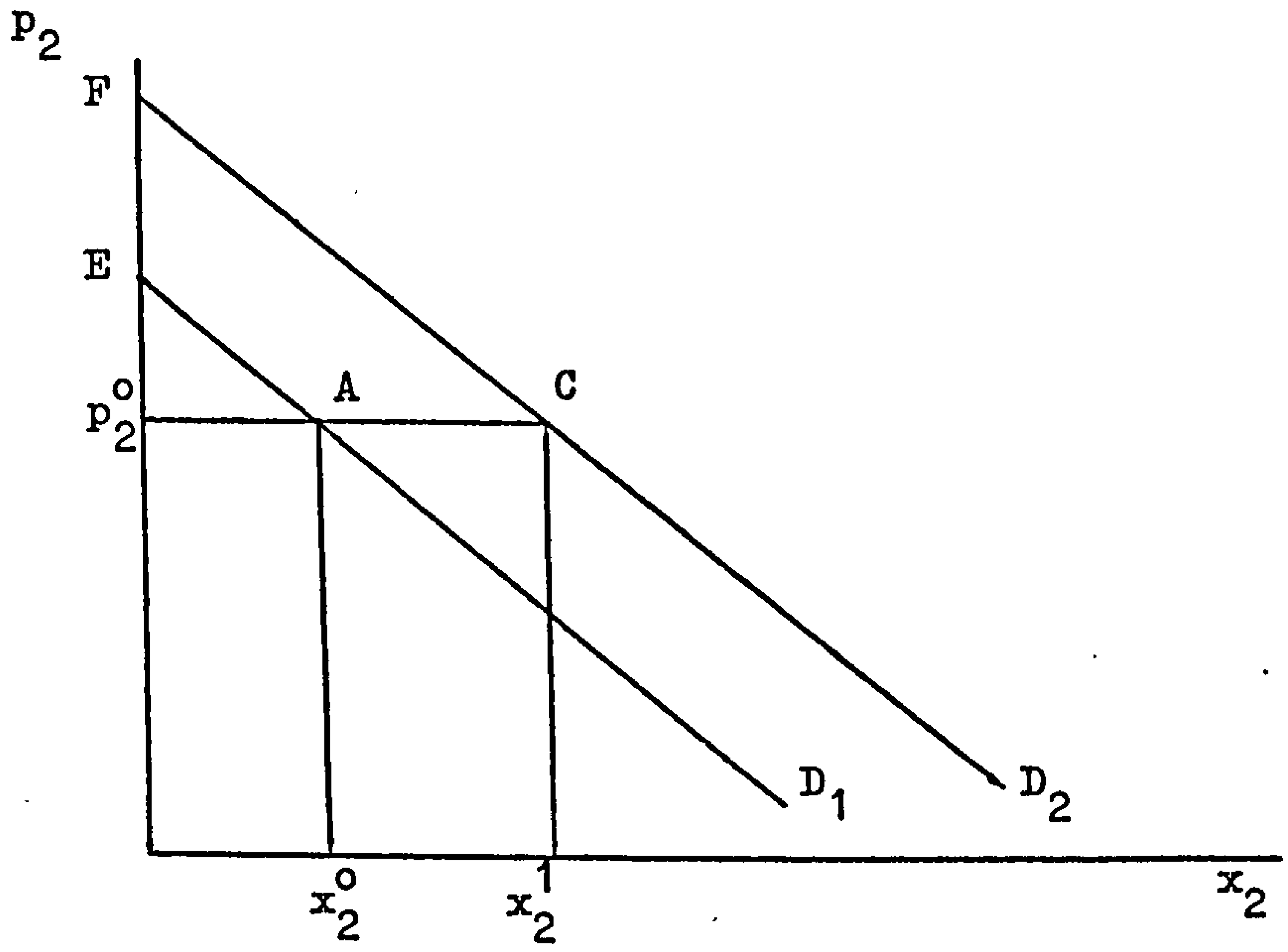
Initially the consumer has indifference curves I^1, I^2, I^3. Following an advertising campaign by the producer of the second good, the consumer's preferences change to I^{11}, I^{22}, I^{33}.

This time it is assumed the consumer knows with certainty the prices of the two goods.

With the original set of indifference curves, the consumer can vary the price of x_2 to again obtain a demand curve for x_2. The initial demand curve D_1 will be fairly close to the origin reflec ing the consumer's preferences for x_1.

Following the advertising campaign the demand curve shifts out, to D_2.

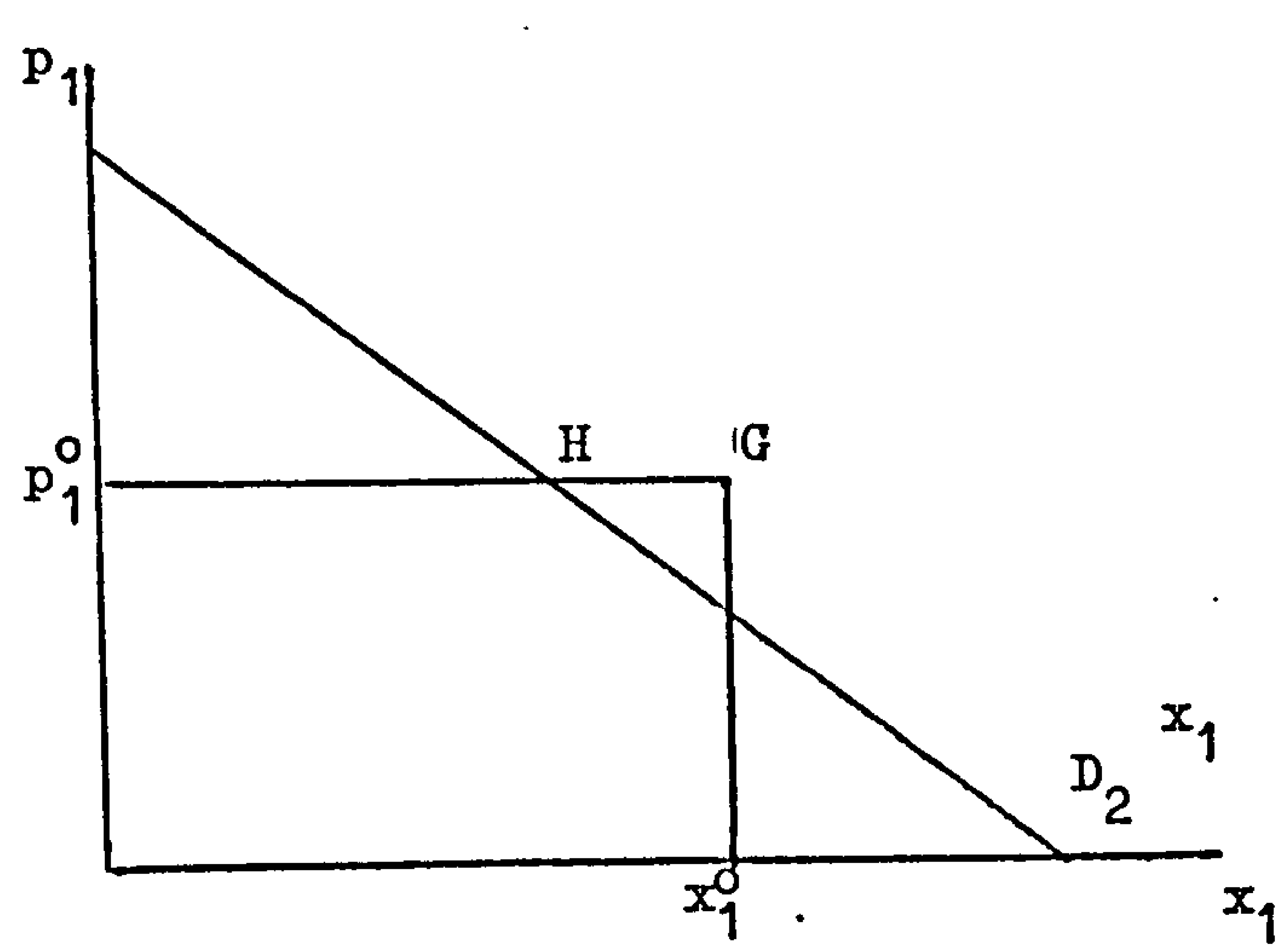
If the consumer is initially at point A , then his consumer surplus is given by Δp_2^o AE.



Following the advertising campaign the consumer moves to point C, and reaps consumer surplus of Δp_2^o CF.

Obviously the consumer surplus is larger in the latter case but is the consumer better off?

For instance, with a given budget constraint if the consumer is demanding more of x_2, he must be demanding less of x_1. Of course with informative advertising, an increase in the demand for x_2 would have meant a fall in the demand for x_1, but this would not have resulted in a fall in welfare, since the consumer would not have been on the true demand curve for x_1 in the original situation.



Initially, with the price expectation for x_2 of $\bar{p}$, the consumer purchased x_1^o at price p_1^o, but at G, the marginal valuation of the good is less than its price; if the consumer moves to point H, his consumer surplus does not change.

But in the case of a shift in preferences, consumer surplus will only increase at the expense of a fall in the consumer surplus of some other good. It could be argued that if all goods advertised, then the demand for all goods would be increased : however this would only be at the expense of leisure. This is essentially Galbraith's argument in the New Industrial State. He argues that in primitive societies, man had a threshold level of consumption, and would work to achieve this threshold and then spend his remaining time on leisure. In modern day capitalist economies this threshold is manipulated upwards by the creation of wants engineered by advertising in order to generate a greater propensity to work from the consumer.

Although Galbraith is credited with this view, Braithwaite had made the same point sometime earlier, "We may, however, imagine the case in which all commodities are advertised so successfully that all subjective valuations are increased. Obviously they cannot all be increased with regard to each other or with regard to money, for this would have no meaning. But they might all be increased with regard to leisure - that is to say, the balance between the marginal disutility of effort and the marginal utility of commodities might be so altered that people would be induced to put forth rather more effort in order to obtain more commodities .... [A]dvertisemen by continually bringing to people's notice the uses and merits of a number of commodities, has so increased their wants that the balance between the disutility of effort and the utility of goods has been altered."(17)

A Special Utility Function

If advertising does change tastes then we are unable to say whether this is good or bad thing. We are not able to state a positive opinion either way. However this does not mean that the analysis ends there. We can side-step the problem by supposing that utility depends upon something that does not change over time. We go back to Lancaster's idea of characteristics. Utility is a function of characteristics: consumers have a preference ordering over characteristics which enables them to construct a set of indifference curves between bundles of characteristics. This preference ordering is not allowed to change over time. These characteristics are produced by consuming goods, the relationship between characteristics and goods is not always known and is allowed to change over time.

The purpose of advertising is to inform consumers about the relationship between a good and its characteristics at a point in time.

We write utility as a function of characteristics

$$U = U(c_1, c_2) \tag{2.1}$$

where c_1 and c_2 are the characteristics; and further, the consumption technology is specified by

$$\left.\begin{aligned} c_1 &= \mu x_1 \\ c_2 &= x_2 \end{aligned}\right\} \tag{2.2}$$

Substituting(2.2)into (2.1)

$$U = U(\mu x_1, x_2) \tag{2.3}$$

Equation(2.3)states utility as a function of goods and the consumption technology matrix represented by μ. We suppose that the value of μ is provided for the consumer by the firm: the firm tells the consumer the value of the elements in the consumption technology matrix. It is assumed that the consumer accepts this value as the true value.

The consumer maximises(2.1)subject to a budget constraint

$$\pi_1 c_1 + \pi_2 c_2 = M$$

where $\pi_1 = \frac{p_1}{\mu}$ and $\pi_2 = p_2$, the π_i are implicit prices for the characteristics. We can easily obtain demand equations for the characteristics as functions of their implicit prices.

Thus in the absence of an advertising message the consumer guesses at the value of the elements in the consumption technology matrix. In fact, the consumer may have a subjective density function describing the possible values of these elements, but let us imagine the consumer makes a "spot" estimate. For this value of the element, by varying the value of p_1 we obtain the perceived demend curve for the consumer; given the actual value of p_1, the consumer locates himself on his perceived demand curve, and we calculate his consumer surplus.

We compare this level of consumer surplus with the consumer surplus which results from locating on a demand curve after the advertising message has been observed. Both of these levels of consumer surplus can then be compared with the level that obtains from the true value of μ. We can thus compare the differences in welfare from making a decision with no advertising assistance and with the help of advertising messages.

For this restricted form of the utility function we will now show how a change in the parameter μ, which is determined by an advertising message, will affect the demand for the good.

The consumer maximises (2.3) subject to a budget constraint

$$p_1 x_1 + p_2 x_2 = M$$

First order conditions yield:

$$\begin{aligned} -\lambda p_1 + U_1 \mu &= 0 \\ -\lambda p_2 + U_2 &= 0 \end{aligned} \qquad (2.4)$$

where λ is the shadow price of the constraint.

Differentiating the first order conditions and the budget constraint with respect to μ; substituting $p_1 = \frac{\mu U_1}{\lambda}$ and $p_2 = \frac{U_2}{p_2}$, and arranging in matrix form:

$$\begin{bmatrix} 0 & \mu U_1 & U_2 \\ \mu U_1 & \mu^2 U_{11} & U_{12} \\ U_2 & \mu U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} -\frac{1}{\lambda}\frac{d\lambda}{d\mu} \\ \frac{dx_1}{d\mu} \\ \frac{dx_2}{d\mu} \end{bmatrix} = \begin{bmatrix} 0 \\ -(U_1 + \mu x_1 U_{11} \\ -x_1 U_{21} \end{bmatrix}$$

If $|U|$ is the determinant of the matrix, then by Cramer's rule:

$$\frac{dx_1}{d\mu} = -\frac{(U_1 + \mu x_1 \quad U_{11})}{|U|} \begin{vmatrix} 0 & U_2 \\ U_2 & U_{22} \end{vmatrix} + \frac{x_1 U_{21}}{|U|} \begin{vmatrix} 0 & U_2 \\ \mu U_1 & \mu U_{12} \end{vmatrix}$$

$$\therefore \frac{dx_1}{d\mu} = \frac{1}{|U|} \{ U_1 U_2^2 + \mu x_1 U_{11} U_2^2 - \mu x_1 U_1 U_2 U_{21} \} \qquad (2.5)$$

which is equivalent to the result obtained by Tintner (1952) [equation (9)] for a two good model with the utility function specified in(2.3). The advantage of this form of the utility function is that marginal utility is not affected by the change in μ.

This same result can be obtained in a less direct but more useful formulation, by specifying the budget constraint in terms of characteristics and a set of implicit prices. That is, we look at the effect of a change in the advertising parameter on the demand for characteristics (in the same way as we look at the effect of a change in price on the demand for a good). We then go on to examine the effect of a change in the demand for a characteristic on the demand for goods.

The consumer maximises (2.1) subject to

$$\pi_1 c_1 + p_2 c_2 = M$$

where π_1 is the implicit price for characteristic 1. $\pi_1 = \frac{p_1}{\mu}$.

Differentiating the first order conditions w.r.t. μ, and arranging in matrix form

$$\begin{bmatrix} 0 & U_1 & U_2 \\ U_1 & U_{11} & U_{12} \\ U_2 & U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} -\frac{1}{\lambda}\frac{d\lambda}{d\mu} \\ \frac{dc_1}{d\mu} \\ \frac{dc_2}{d\mu} \end{bmatrix} = \begin{bmatrix} \frac{U_1 c_1}{\mu} \\ -\frac{U_1}{\mu} \\ 0 \end{bmatrix}$$

$$\frac{dc_1}{d\mu} = -\frac{U_1 c_1}{\mu |D|} \begin{vmatrix} U_1 & U_{12} \\ U_2 & U_{22} \end{vmatrix} - \frac{U_1}{\mu} \frac{1}{|D|} \begin{vmatrix} 0 & U_2 \\ U_2 & U_{22} \end{vmatrix} \qquad (2.6)$$

where $|D|$ is the determinant of the matrix and if $c_1 = \mu x_1$

Then $\dfrac{dc_1}{d\mu} = \dfrac{\mu dx_1}{d\mu} + x_1$ (2.7)

and substituting equation (2.6)

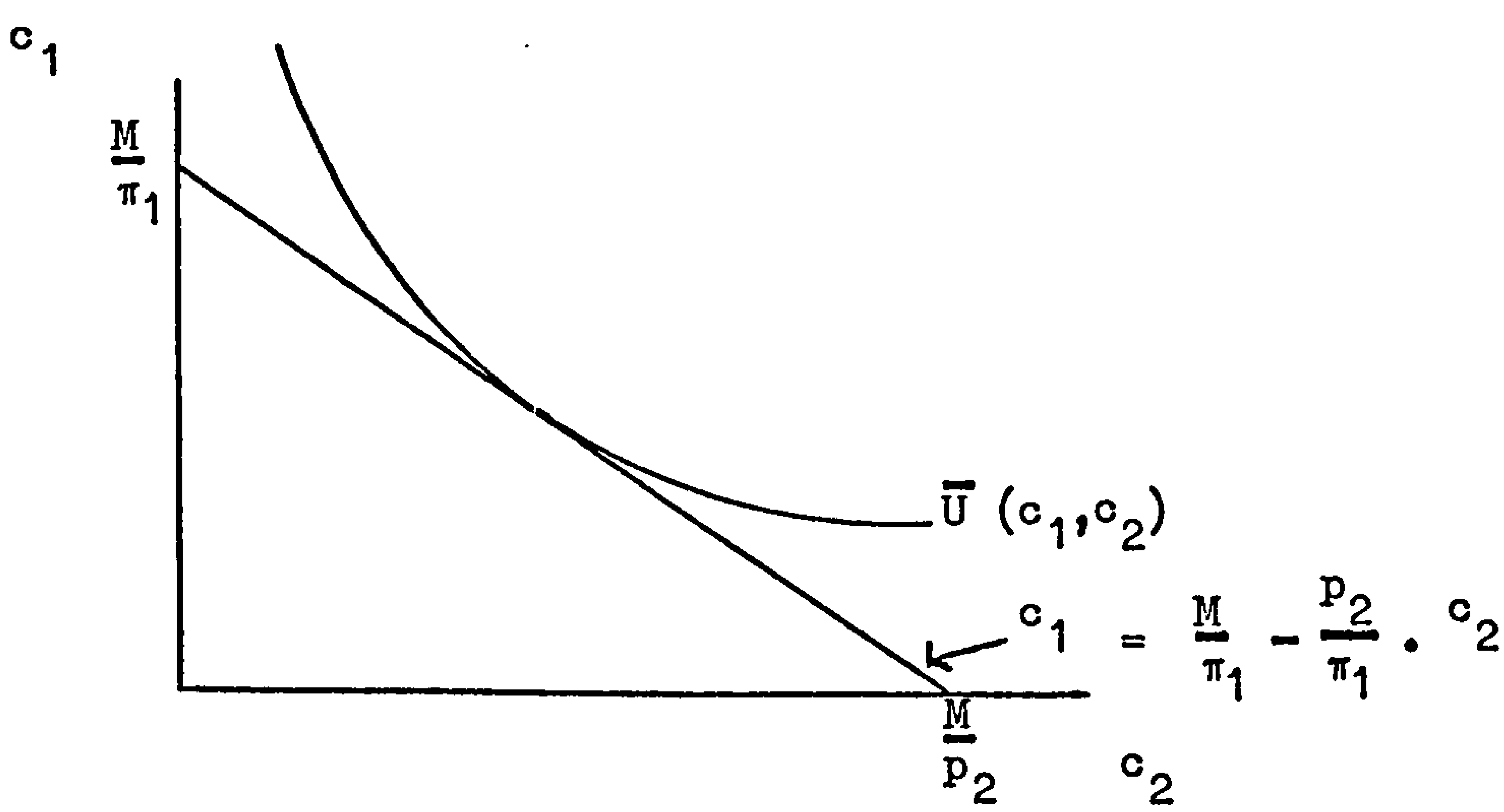
$$\therefore \frac{dx_1}{d\mu} = \frac{1}{a^2 \; |D|} \{ - \mu x_1 U_1^2 U_{22} + \mu x_1 U_1 U_2 U_{12} + U_1 U_2^2 - \mu x_1 |D|$$

But $a^2 |D| = |U|$

$$\therefore \frac{dx_1}{d\mu} = \frac{1}{|U|} \{ U_1 U_2^2 + \mu x_1 U_{11} U_2^2 - \mu x_1 U_1 U_2 U_{21} \} \qquad (2.8)$$

and (2.8) is the same result as (2.5).

The advantage of this less direct route is that we are able to look at equation (2.6) in terms of income and substitution effects



By changing the parameter μ, we change the implicit price π_1, and consequently the slope of the budget constraint.

Consider the implicit price effect. Differentiate the first order conditions w.r.t. π_1

$$\begin{bmatrix} 0 & U_1 & U_2 \\ U_1 & U_{11} & U_{12} \\ U_2 & U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} -\frac{1}{\lambda} \frac{d\lambda}{d\pi_1} \\ \frac{dc_1}{d\pi_1} \\ \frac{dc_2}{d\pi_2} \end{bmatrix} = \begin{bmatrix} -c_1\lambda \\ \lambda \\ 0 \end{bmatrix}$$

$$\frac{dc_1}{d\pi_1} = \frac{c_1\lambda}{|D|} \begin{vmatrix} U_1 & U_{12} \\ U_2 & U_{22} \end{vmatrix} + \frac{\lambda}{|D|} \begin{vmatrix} 0 & U_2 \\ U_2 & U_{22} \end{vmatrix} \qquad (2.9)$$

where the first term in(2.9)is the income effect and the second term is the substitution effect.

Making the substitution $\lambda = \frac{\mu U_1}{p_1}$, we may write(2.9)as

$$\frac{dc_1}{d\pi_1} = \frac{\mu U_1}{p_1} \frac{c_1}{|D|} \begin{vmatrix} U_1 & U_{12} \\ U_2 & U_{22} \end{vmatrix} + \frac{\mu U_1}{p_1} \cdot \frac{1}{|D|} \begin{vmatrix} 0 & U_2 \\ U_2 & U_{22} \end{vmatrix} \qquad (2.10)$$

We can compare(2.10)with (2.6).

$$\frac{dc_1}{d\mu} = \frac{dc_1}{d\pi_1} \cdot \frac{d\pi_1}{d\mu} = -\frac{dc_1}{d\pi_1} \cdot \frac{p_1}{\mu^2}$$

Multiplying(2.10)through by $-\frac{p_1}{\mu^2}$ we obtain (2.6).

Thus from(2.7),we may write

$$\frac{dx_1}{d\mu} = \frac{1}{\mu}\left[-\frac{dc_1}{d\pi_1}\cdot\frac{p_1}{\mu^2}\right] - \frac{x_1}{\mu} . \qquad (2.11)$$

The effect of an increase in μ - a message informing the consumer that the amount of c , per unit of x_1 has increased - on the demand for x_1 , is made up of two separate effects. Firstly, an increase in μ will increase the demand for c_1 , and the impact of the increased demand for characteristics will be tempered by the amount of c_1 provided by one unit of x_1 . The second effect occurs because some of the increased desire for c_1 will be satisfied by the amount of x_1 already consumed, since an increase in μ means that the consumer can now purchase the same quantity of the characteristic by buying less of the good, or at least, believes he can do this. Thus the change in the demand for the good caused by an increase in the advertised level of product quality may be positive or negative

The results we have obtained concerning the possibility of welfare comparisons and looking at the effect of an advertising message on the demand for a good apply to a specific utility function. However, the analysis can be made more general: whatever the form of the original utility function it is only necessary to redefine it in terms of variables that do not change over time, and switch the impact of advertising from the utility surface to an implicit price line.

Multiplying(2.10)through by $-\frac{p_1}{\mu^2}$ we obtain (2.6).

Thus from(2.7),we may write

$$\frac{dx_1}{d\mu} = \frac{1}{\mu}\left[-\frac{dc_1}{d\pi_1}\cdot\frac{p_1}{\mu^2}\right] - \frac{x_1}{\mu}\,. \tag{2.11}$$

The effect of an increase in μ - a message informing the consumer that the amount of c, per unit of x_1 has increased - on the demand for x_1, is made up of two separate effects. Firstly, an increase in μ will increase the demand for c_1, and the impact of the increased demand for characteristics will be tempered by the amount of c_1 provided by one unit of x_1. The second effect occurs because some of the increased desire for c_1 will be satisfied by the amount of x_1 already consumed, since an increase in μ means that the consumer can now purchase the same quantity of the characteristic by buying less of the good, or at least, believes he can do this. Thus the change in the demand for the good caused by an increase in the advertised level of product quality may be positive or negative

The results we have obtained concerning the possibility of welfare comparisons and looking at the effect of an advertising message on the demand for a good apply to a specific utility function. However, the analysis can be made more general: whatever the form of the original utility function it is only necessary to redefine it in terms of variables that do not change over time, and switch the impact of advertising from the utility surface to an implicit price line.

Conclusions

The importance of distinguishing between informative and persuasive advertising will vary depending upon the definitions given to these two types of advertising. We have argued that all advertising is informative; an advertising message is a statement with an information content, which may need to be presented in a persuasive way.

As Michan (1969) comments "the terms information and persuasion, even in their purest forms, are not necessarily opposed in meaning"[18]. If a piece of information is to reach the consumer, to allow him to make an informed decision it is necessary that the information be presented persuasively, similarly "the persuasive potential of an advertisement .... may well be increased with the amount of information provided".[19] For Mishan the critical issue is not whether advertising is informative or persuasive but whether the information presented is impartial or biased. Mishan maintains that advertising messages will consist of biased messages, which may include information which is not factually correct. These biased messages may also be presented in such a way as to prevent objective evaluation since the messages are deliberately vague and can not be substantiated. We have argued that the welfare effects of this type of advertising can be calculated. Advertising messages may also provide information about parameters in the utility function. If the utility function can be transformed to one in which the advertising statements affect the implicit prices of the redefined variables and not the utility function, then again we can make meaningful welfare comparisons. Advertising which provides information about new characteristics can not be coped with. One way out of this impasse is to simply suppose that there are no such things as new characteristics. Utility is a function of a finite number of known sensations which can be achieved by a large number of goods, some of which are as yet undiscovered. Of course it may be felt that this assumption is unrealistic. In which case an alternative solution is to accept that advertising which provides information about new tastes can not be judged in terms of welfare improvements. We simply acknowledge that positive economics can not always say whether this kind of advertising is a good or bad thing.

Some writers have challenged the usefulness of making a distinction between informative and persuasive advertising. Fulop (1981) summarises the 'Austrian' viewpoint

"The artificial distinction drawn between informative and persuasive advertising is based on a misunderstanding of the nature, purpose and rationale of advertising which is wider than conveying information to the consumer about products which are already in existence (new products; continually changing consumers - baby products; new uses of an established product; changes in the product itself)."[20]

In the concluding chapter of their book, Chiplin and Sturgess argue

"While we consider that the dichotomy between informative and persuasive advertising has no operational significance, it remains true that advertising can be a substantial influence in shaping consumer preferences. Economists have no clear criteria for assessing the consequences of advertising in a world of changing tastes."[21]

But the fact that we cannot make welfare comparisons when tastes change but we can make welfare comparisons when advertising provides information, truthful or not, means that the distinction is operationally significant!

For a real world example of the policy difficulties distinguishing between informative and persuasive advertising, consider the case of professional opticians who are not allowed to advertise the price of their product at present. Without advertising consumers operate with imperfect information, and a consequence of this is that opticians are able to charge different prices for the same good at different locations. Unless consumers engage in costly search, they will pay the first price encountered. It is argued that advertising would end this practice, of localised monopolies, since if the public were made aware of the prices being charged, the competitive process would ensure one industry price which just covered average costs and normal profits.

However, although we can see that this type of advertising will increase consumer welfare, once advertising is allowed there is no reason to believe that it will only entail advertising about price. It is also likely that opticians will either individually or collusively broadcast the virtues of wearing spectacles and attempt to increase the demand for their good by informing consumers about the 'unrealised' benefits of spectacles, in a persuasive way. That is, they attempt to change tastes. From a policy point of view we may advise that advertising about prices which will assist the competitive process will be a good thing; but if advertising shifts preferences we are unable to make a positive statement.

To conclude, we have argued that the important aspect of advertising is the information it imparts, since the consumer will act upon receiving this information whether it is true or false. We have seen that for particular types of utility functions we can make positive statements about welfare improvements. We have examined the effect of an advertising message which states that product quality has improved, on the demand for that product.

Notes

(1) Marshall (1920) p. 304.
(2) Marshall (1920) p. 306
(3) Marshall (1920) pp. 306-7
(4) Marshall (1920) pp. 304-5
(5) Pigou (1929) p. 198
(6) Bain (1956) p. 114
(7) Joyce (1963) p. 123
(8) Joyce (1967) p. 152
(9) Hicks (1962) p. 257
(10) Marshall (1920) p. 76
(11) Pigou (1929) p. 198
(12) Marshall (1920) p. 306
(13) Marshall (1920) p. 306 footnote
(14) Braithwaite (1928) p. 17
(15) Braithwaite (1928) pp. 19-21
(16) Braithwaite (1928) pp. 18-19
(17) Braithwaite (1928) p. 20
(18) Mishan (1969) p. 114
(19) Mishan (1969) p. 115
(20) Fulop (1981) p. 14
(21) Chiplin and Sturgess (1981) p. 133

Chapter 3

Advertising Under Imperfect Information

1. Introduction and the Notion of Ignorance

Neoclassical economics is all about choice. Producers choose which goods to produce and what factors to use; consumers choose which goods to buy and how much of a factor to supply. An assumption behind these decision processes is that the economic agents know what they are doing. Consumers are assumed to have sufficient knowledge to be able to construct transitive preference orderings of all the goods available, know the quality of each good, in so much as quality variations are allowed, and be able to rank their preference orderings.

A way of incorporating quality differences between goods is to make use of Lancaster's characteristic approach to demand analysis. Utility is a function of characteristics and characteristics are related to goods through the consumption technology matrix.

$$u = u(c)$$

$$c = A\,x$$

where c is a vector of characteristics, x is a vector of goods and A is the consumption technology matrix which transforms the goods into characteristics.

This less traditional approach is still based on the assumption that people have complete information and in fact introduces an additional tier of knowledge Consumers now have to know about the effect of characteristics on utility, and the relationship between goods and characteristics.

It is important to distinguish at the outset between exogenous and endogenous tastes. Tastes may be innate, inherent from birth or they may be determined within the environment in which the consumer operates. The view taken here is that having utility depend directly upon characteristics rather than goods means that we can suppose the consumer is aware of the basic characteristics that yield him utility. New products are only a rearrangement of existing characteristics. Totally new characteristics are not allowed to enter the consumer's objective function at a later date.

The generalisation that economic agents live in a world of complete knowledge is made up of a number of separate assumptions concerning particular aspects of the environment. Firstly, in taking a decision an agent is aware of the opportunity costs in terms of the set of alternative actions forgone; secondly, the agent knows the consequences of any action he takes; and finally the agent can actually identify the objective function he is attempting to maximise.

If all the above assumptions hold then it can be said that the world has perfect information. But if only some or none of the above hold then the world is far from perfect and in order to make decisions agents must act with less than complete knowledge or make a positive effort to gain additional information. The process of gaining information is unlikely to be costless and different methods of generating information will have differing costs depending upon the type of information being sought. The type of information being sought will depend upon the type of information that is absent in the agents decision problem. This chapter is concerned with establishing advertising as an example of an information generating technique that is available to an agent, and examining the relationship between advertising and the nature of the information imperfection.

A neoclassical agent maximises an objective function by choice of his decision variables. Four types of ignorance can be identified: (a) ignorance of the objective function; (b) ignorance of the elements in the set of feasible decisions; (c) ignorance of the effect a decision will have on the objective function; and (d) in the presence of a stochastic variable, ignorance of the objective probability distribution. The agent may face one or more types of ignorance. As an illustration of the first three, consider a consumer who firstly does not know the values of the parameters in his utility function, that is, he is unsure as to what he is attempting to maximise. Secondly he is unaware of the range of substitute goods, and finally he does not know the quality of the product, and in purchasing a good does not know the utility that will result from it.

Turning to the fourth type of ignorance, instead of the hitherto implicit assumption that the world is deterministic, let a stochastic variation be incorporated, then agents are assumed to have distribution functions over the uncertainty, and the parameters of these distribution functions may be unknown. For example, if the pay-off from a good is stochastic, say the utility from an umbrella depends upon the weather, then the consumer may not know the probability of a particular state of weather occurring, which is necessary to calculate the expected utility from an umbrella.

We have listed the occasions when imperfect knowledge arises in the consumer's decision problem of choosing a bundle of goods to maximise utility. As the nature of the imperfection varies, the way that the consumer overcomes his ignorance and its cost will also vary. Given that the consumer is in possession of only imperfect information, he has a choice: he can go ahead and make a decision on the basis of the limited amount of information available or he can incur a cost, obtain additional information and then make his decision. The initial decision to obtain further information or not, will depend on the costs relative to the benefits.

In the next section of this chapter we consider the expected value of both perfect and imperfect information. In Section 3 we illustrate the expected and actual values of perfect information with an example. In Section 4, we introduce the role of advertising and extend the example from the previous section. We develop a model of how advertising messages are incorporated in a consumer's beliefs, and contrast advertising with other sources of information. Section 5 is concerned with the arguments about the information content of advertising statements. The final section summarises the main points.

2. Value of Information

Imperfect information can be modelled by assuming that the unknown parameter, wherever it occurs, is a random variable to the consumer, and as such has an associated subjective distribution function. Gould (1974) shows that the value of perfect information is non-negative.

If z is an unknown parameter and x is the decision variable, the decision maker receives a pay-off of $U(z, x)$ if z obtains when he has chosen x.

Suppose all the information available is represented by the distribution function $F(z)$, then the consumer chooses x to maximise $EU(z, x)$:

$$EU = \int_{\forall z} U(z, x)\, dF(z) \tag{3.1}$$

Let x^* be the value of x which maximises this expression.

Alternatively, suppose the consumer could pay a fee and observe the unknown parameter before making a decision. In this case the consumer will know the value of z before choosing x, and will choose x to maximise $U(z, x)$. Let the value of this x, be x_z^*. Before receiving the information the consumer expects a pay-off of

$$\int U(z, x_z^*)\, dF(z) \tag{3.2}$$

Marschak (1954) gives as an example a gambler who pays 50p to bet on whether a coin comes up heads or tails; if he gets it right he wins £1. Obviously his expected gain is zero. This can be contrasted with a situation where a gambler pays someone to observe whether the coin has come up heads or tails before the bet is placed. The gambler would pay up to 49p for this information, since his expected gain is 50p. By definition $U(z, x_z^*) \geq U(z, x)\ \forall\ x$ and since $F(z)$ is non-decreasing in z

$$\int U(z, x_z^*)\, dF(z) \geq \int U(z, x)\, dF(z)$$

$$\therefore \int U(z, x_z^*)\, dF(z) \geq \max_x \int U(z, x)\, dF(z) = \int U(z, x^*)\, dF(z)$$

Thus the value of perfect information is non-negative

$$\int U(z, x_z^*)\, dF(z) - \int U(z, x^*)\, dF(z) \geq 0 \tag{3.3}$$

since it is the expected value of utility if the random variable is known less the expected value of utility if the parameter is unknown. If the expected value of perfect information is greater than its costs, we can expect that the consumer would choose to obtain further information.

It can be seen that the imperfectly informed consumer has a choice between making a decision based on the subjective distribution function of the unknown parameter, or incur a cost and obtain perfect information before continuing with the decision process. The act of a consumer paying for and obtaining extra information can be termed search, whereas a consumer who chooses not to search, makes a random sample.

Nelson (1970) draws a distinction between search goods and experience goods. The quality of search goods can be ascertained prior to purchase, but the quality of experience goods can only be determined after the good has been bought. For example, a dress can be examined before purchase and is a search good; a can of tuna fish must be bought before it can be opened and tested. To an extent it can be visualised that all goods can be searched prior to purchase, but in the case of some goods the costs of search are prohibitive, and these goods are labelled experience goods. Thus in reality, there is a spectrum of goods, ranging between the two extremes of search and experience and the location of the good along the spectrum will depend upon the cost to the consumer of ascertaining its quality.

The distinction between search and experience goods fits neatly into the notion of the expected value of information. Suppose the unknown parameter, z, represents product quality. When the expected value of information about the quality of a good is greater than its search cost, which may in some cases be zero, then the good is a search good. If the cost of search is greater than the expected value of information then the good is an experience good. In the case of an experience good it does not pay the consumer to find out about produc quality before purchasing the good, and he will make a decision solely on the basis of the subjective distribution function.

Uncertainty about product quality can be extended to a more general treatment by making use of Lancaster's approach, where any characteristic can be obtained through more than one good, and different goods contain characteristics in differing proportions. The relationship is described by the consumption technology matrix, and uncertainty about product quality can be represented by assuming that the elements in the A-matrix are unknown. So the consumer will have subjective distributions over each of the unknown parameters in the A-matrix. We have considered the expected value of information in relation to product quality, but can this same concept be applied to the other types of ignorance: ignorance of parameters in utility function and ignorance of substitute goods? It is easy to incorporate imperfect knowledge of the utility function into the framework, by letting ε be any parameter in the utility function, not just product quality. Reference to equations (3.1) and (3.2) show that this more general case has already been allowed for. In this case the consumer is able to pay a search cost in order to find out the effect a characteristic will have on utility.

The third type of ignorance: ignorance of competing products can be fitted into the model by supposing that the consumer acts as if the elements in the consumption technology matrix are zero.

Suppose that the consumer gains utility from a vector of characteristics with elements c_i $(i=1,\ldots,n)$ and there is a vector of goods available with elements x_j $(j=1,\ldots,m)$. However the consumer only knows of the existance of k goods, where $k < m$. Then the consumption technology matrix consists of the following elements

$$A = \left[\begin{array}{ccc|cc} z_{11} & z_{12} \cdots & z_{1k} & z_{1k+1} \cdots\cdots & z_{1m} \\ z_{21} & z_{22} & z_{2k} & z_{2k+1} & \\ \vdots & \vdots & & & \\ z_{n1} & z_{n2} & z_{nk} & z_{nk+1} & z_{nm} \end{array}\right].$$

The consumer assigns subjective estimates to these unknown parameters. For goods which the consumer is unaware of, he acts as though he believes the parameter is zero. Thus ignorance about competing products is really a special case of ignorance of product quality. The fact that a consumer does not know that a particular good exists is the same as assuming that the consumer knows it exists but does not know its function, or believes it has no function. Imagine that a consumer goes to buy a jar of jam, and alongside all the known brands there is a new jar labelled in a language the consumer does not understand. Does the new jar contain jam or not? The coefficient in the technology matrix between the jar of the new good and the characteristic jam is either zero or one. Either the jar contains jam or it does not. The consumer will have a distribution function over the probability that the new jar is a pot of jam, and will make a decision to buy the new product or not according to these probabilities.

Ignorance of competing products becomes a special case of ignorance of product quality. A consumer has a utility function over a vector of characteristics, achieved by purchasing a certain amount of a certain number of an entire vector of all possible goods. The relationship between goods and characteristics is through the technology matrix composed of elements which are unknown to the consumer and may take on zero or positive values.

Equations (3.1) and (3.2) are general statements of expected utility when the consumer faces a vector of unknown parameters representing different types of imperfect information. Equation (3.3) gives the expected value of information for all these types of ignorance. Having said that, it can be seen that a new type of information has appeared. In order to calculate expected utilities the consumer is assumed to know the parameters in the subjective distribution function. There is no reason to believe that the consumer does know these parameters. This problem can be surmounted by supposing that if the consumer has no idea about the shape of the distribution then it can be approximated by a uniform distribution, The consumer is now in a position of total ignorance, he faces an unknown parameter in the utility function, over which he has a diffuse subjective probability distribution. Actually this situation is not the worst possible for the consumer. Gould asks the question: what is the nature of the distribution function when the expected value of information, equation(3.3), is maximised? He shows that the maximum value of information does not necessarily increase as the number of possible outcomes increases nor does it necessarily involve equiprobable outcomes, i.e. it does not have to be a uniform distribution!

Now suppose a message is made available to the consumer, which could be an advertising statement, giving the consumer information about the unknown parameter. Then this statement has value. If the advertising message was known to provide perfect information then its expected value would be given by(3.3). However the consumer cannot be sure that the message reflects the true value of the parameter, but the consumer's subjective beliefs will be influenced by the advertising message. The conditional subjective distribution of z, given the statement μ, is represented by $G(z|\mu)$, and following the advertising message the consumer maximises:

$$\max_{x} \int U(z, x)\, dG\,(z|\mu) \qquad (3.4)$$

However, the advertising message has still to be observed; the consumer will weight expression(3.4)by the probability of any value of μ occurring. If x^*_μ is the value of x which maximises(3.4), then before receiving the advertising message the consumer expects a pay-off of:

$$\int \left[\int U(z,\, x_\mu^*)\ dG\ (z\,|\,\mu) \right] dJ\ (\mu) \qquad (3.5)$$

where $J\ (\mu)$ is the consumer's subjective distribution of the possible values that the advertising message might take.

We can write the expected value of imperfect information as:

$$\int \left[\int U(z,\, x_\mu^*)\ dG\ (z\,|\,\mu) \right] dJ\ (\mu)\ -\ \int U(z,\, x^*)\ dF\ (z) \qquad (3.6)$$

But the Theorem of Total Probabilities tells us that:[1]

$$\int dG\ (z\,|\,\mu)\ dJ\ (\mu)\ =\ dF\ (z) \qquad (3.7)$$

Thus consistance requires that the consumer's conditional beliefs about the unknown parameter given the advertising message, $G(z\,|\,\mu)$, his beliefs about the possible values that the advertising message might take, $J\ (\mu)$, and his unconditional prior beliefs about z, $F\ (z)$, are related by the constraint in (3.7).

Substituting(3.7)into(3.6),we may write

$$V\ =\ \int \int U(z,\, x_\mu^*)\ dG\ (z\,|\,\mu)\ dJ\ (\mu)\ -\ \int \int U(z,\, x^*)\ dG\ (z\,|\,\mu)\ dJ\ (\mu)$$

$$V\ =\ \int \left[\int \{ U(z,\, x_\mu^*) - U(z,\, x^*) \}\ dG\ (z\,|\,\mu) \right] dJ\ (\mu) \qquad (3.8)$$

But by definition $\int \{U(z,\, x_\mu^*) - U(z,\, x^*)\}\ dG(z\,|\,\mu) \geq 0$ since x_μ^* was picked to maximise.

$$\int u(z,x)\ dG(z\,|\,\mu)$$

$$\therefore\ V\ \geq 0\ .$$

So the expected value of imperfect information is non-negative.[2]

The actual value of information is an ex post value and depends on the content of the information. It is the difference between utility after the message has been observed and a decision taken and the utility that would have resulted from taking a decision which maximised expected utility, when information was absent. If the information is made available at a cost then the actual value will be net of any search cost. In the case of advertising the cost of the message is already included in the price of the product as advertising and the goods are joint products. According to Telser (1966) the economies from this joint production mean that "the total resources engaged in the supply of advertising messages would be less than if the advertising messages were sold separately from the physical goods".[3] Advertising may be a more efficient method of generating information than other techniques.

The ex post value of information indicates by how much utility is higher because consumers took a fully informed decision as opposed to the utility from a decision taken in ignorance. The distinction between the expected value of information and the actual value can be brought out by an example.

3. An example

Consider a consumer who has trudged through the desert to the supermarket at the oasis only to find that none of the commodities on the shelves are labelled. He is faced with a mass of homogeneously packed goods, unable to tell the difference between cartons of chalk and cheese. Not surprisingly after the long trek he is thirsty and required a drink, but which package should he choose? In his dehydrated state he is unlikely to gain utility from any good which does not contain some kind of drink. His utility over characteristics can be expressed as a function of drink alone: $u = f(c) = \sqrt{c}$ where c is the characteristic drink. The consumer brings with him an amount of money, M, and is able to narrow the choice of goods yielding utility down to two, x_1 and x_2. But he is not sure which good contains drink and which does not. If he purchases x_1, of the first good, he received a pay-off of $\alpha_1 x_1$ where α_1 is an unknown parameter which can take on the values 0 or 1. Let $(1 - \Pi)$ be the probability that α_1 takes on the value of 1. If the consumer purchases x_2 of the second good, he receives $\alpha_2 x_2$ and if the events are mutually exclusive, Π is the probability that $\alpha_2 = 1$. The consumer maximises expected utility:

$$Eu = (1 - \Pi)\, x_1^{\frac{1}{2}} + \Pi\, x_2^{\frac{1}{2}}$$

$$\text{s.t.} \quad x_1 + p\, x_2 = M$$

where p is the relative price of good 2 to good 1.

First order conditions yield:

$$x_2 = \frac{M}{p} \cdot \frac{\Pi^2}{p(1-\Pi)^2 + \Pi^2}$$

$$x_1 = M \cdot \frac{p(1-\Pi)^2}{p(1-\Pi)^2 + \Pi^2} \tag{3.9}$$

Substituting back into the expected utility function gives the value of the maximised expected utility:

$$Eu = \left(\frac{M}{p}\right)^{\frac{1}{2}} \left[(1 - \Pi)^2 p + \Pi^2\right]^{\frac{1}{2}} \tag{3.10}$$

If instead the consumer had been told which package contained which good, he could have spent his entire income on the package containing drink. In this case his expected utility before receiving the information is:

$$\begin{aligned} Eu^* &= (1 - \Pi)M^{\frac{1}{2}} + \Pi\left(\frac{M}{p}\right)^{\frac{1}{2}} \qquad (3.11) \\ &= \left(\frac{M}{p}\right)^{\frac{1}{2}} \left[p^{\frac{1}{2}} (1 - \Pi) + \Pi\right] \end{aligned}$$

The expected value of information is Eu* - Eu

$$= \left(\frac{M}{p}\right)^{\frac{1}{2}} \left[p^{\frac{1}{2}} (1 - \Pi) + \Pi - \left[(1 - \Pi)^2 p + \Pi^2\right]^{\frac{1}{2}}\right] \tag{3.12}$$

<u>Proposition</u> To show equation(3.12)is non-negative.

<u>Proof</u> (3.12)is non-negative if

$$p^{\frac{1}{2}} (1 - \Pi) + \Pi \quad > \{(1 - \Pi)^2 p + \Pi^2\}^{\frac{1}{2}}$$

Square both sides

$$p(1 - \Pi)^2 + \Pi^2 + 2p^{\frac{1}{2}}\Pi(1 - \Pi) > (1 - \Pi)^2 p + \Pi^2$$

i.e. if $\qquad 2p^{\frac{1}{2}} \Pi(1 - \Pi) > 0$

Q.E.D.

The actual value of information depends upon the realised values of the unknown parameters. If $\alpha_1 = 1$ and $\alpha_2 = 0$, then a consumer who maximised expected utility would have purchased a mixture of the two goods as shown by equation(3.4 This would yield a utility level in this case of:

$$\left[\frac{M\, p(1 - \Pi)^2}{p(1 - \Pi)^2 + \Pi^2}\right]^{\frac{1}{2}}$$

as opposed to a utility of $M^{\frac{1}{2}}$ if the consumer had known that $\alpha_1 = 1$. The actual value of information is

$$M^{\frac{1}{2}} \left[1 - \frac{p(1 - \Pi)^2}{p(1 - \Pi)^2 + \Pi^2} \right]^{\frac{1}{2}} \qquad (3.13)$$

If $\alpha_1 = 0$ and $\alpha_2 = 1$, then maximising expected utility yields an actual utilit of

$$\left[\frac{M}{p} \cdot \frac{\Pi^2}{p(1 - \Pi)^2 + \Pi^2} \right]^{\frac{1}{2}}$$

as opposed to an utility of $(\frac{M}{p})^{\frac{1}{2}}$ if the consumer had known that $\alpha_2 = 1$. The actual value of information in this case is

$$(\frac{M}{p})^{\frac{1}{2}} \left[1 - \frac{\Pi^2}{p(1 - \Pi)^2 + \Pi^2} \right]^{\frac{1}{2}} > 0 \qquad (3.14)$$

In both cases, the actual value of information is non-negative. An accurate advertising statement which informs the consumer about α_1 and α_2 has an actual value given by equation(3.13)or(3.14). These expressions show the extra utility gained from consumers making a fully informed decision as opposed to the utility gained from a decision taken in an ignorant state.

The Role of Advertising

Now suppose the consumer is able to observe an advertising message μ before making a decision. But advertising is not straightforward information, it has a peculiar nature in that it is provided by the person who wishes to sell the product. For example if $\alpha_2 = 1$ and $\alpha_1 = 0$, then the producer of the first good who knows that the consumer only gets utility from drink, and will not purchase good 1 if he knows it is not a drink, has an incentive to advertise that $\alpha_1 = 1$. The consumer must decide whether an advertisement is telling the truth or not.

The consumer is aware that an advertiser may not provide accurate information and based on his beliefs about the truthfulness of the message will discount the advertising statement to some degree. We suppose that the consumer does not accept the advertising message as perfect information, but does use it to update his beliefs about the unknown parameter. The consumer holds his beliefs about α_1 and α_2 conditional upon the possible values of μ. Let the possible values for μ be either $(0, 1)$. Then the consumer has four conditional probabilities:

$$\Pr[\alpha_1 = 1 \mid \mu = 1] \qquad \Pr[\alpha_1 = 0 \mid \mu = 1]$$

$$\Pr[\alpha_1 = 1 \mid \mu = 0] \qquad \Pr[\alpha_1 = 0 \mid \mu = 0]$$

Let μ be supplied by the producer of x_1, then it seems reasonable to suppose that if this producer does not advertise, i.e. $\mu = 0$, it is certainly because his good does not contain the characteristic desired by our thirsty consumer.

Thus $Pr[\alpha_1 = 1 \mid \mu = 0] = 0$ and $Pr[\alpha_1 = 0 \mid \mu = 0] = 1$ (3.15)

On the other hand the fact that the producer has advertised, i.e. $\mu = 1$, does not ensure that his good contains the characteristic, and the consumer's beliefs given the advertising message are as follows:

$$Pr[\alpha_1 = 1 \mid \mu = 1] = 1 - \Pi_\mu \; ; \qquad Pr[\alpha_1 = 0 \mid \mu = 1] = \Pi_\mu \qquad (3.16)$$

The consumer will use these conditional beliefs described in equations (3.15) and (3.16) to maximise expected utility

$$Eu = (1 - \Pi_\mu)\, x_1^{\frac{1}{2}} + \Pi_\mu x_2^{\frac{1}{2}} \qquad \text{if } \mu = 1$$

$$Eu = x_2^{\frac{1}{2}} \qquad \text{if } \mu = 0$$

$$\text{s.t. } x_1 + px_2 = M$$

Substituting back the optimal values of x_1 and x_2 into the expected utility functions, gives the value of the maximised expected utility.

$$Eu = \left(\frac{M}{p}\right)^{\frac{1}{2}} \left[(1 - \Pi_\mu)^2 p + \Pi_\mu^{\,2} \right]^{\frac{1}{2}} \qquad \text{if } \mu = 1 \qquad (3.17)$$

$$EU = \left(\frac{M}{p}\right)^{\frac{1}{2}} \qquad \text{if } \mu = 0 \qquad (3.18)$$

To calculate the expected utility before observing the value of μ involves weighting the expressions in (3.17) and (3.18) by the probability of μ taking on either value. Let $Pr[\mu = 1] = \omega$ and $Pr[\mu = 0] = 1 - \omega$.

Then the expected utility of the consumer before observing the advertising message is

$$Eu^{**} = (\frac{M}{p})^{\frac{1}{2}} \cdot [(1 - \Pi_\mu)^2 p + \Pi_\mu{}^2]^{\frac{1}{2}} \omega + (\frac{M}{p})^{\frac{1}{2}} (1 - \omega) \qquad (3.19)$$

As was noted earlier the conditional beliefs about α , the prior beliefs about μ and the prior beliefs about α are related by the constraint in(3.7). For our example with discrete variables this constraint reduces to

$$(1 - \Pi) = Pr [\alpha_1 = 1] = \sum_{i=0}^{1} Pr [\alpha_1 = 1 \mid \mu = i] Pr [\mu = i]$$

$$= 0 \cdot (1 - \omega) + (1 - \Pi_\mu) \omega$$

$$\therefore (1 - \Pi) = (1 - \Pi_\mu) \omega \qquad (3.20)$$

and $$\Pi = Pr [\alpha_2 = 1] = \sum_{i=0}^{1} Pr[\alpha_1 = 0 \mid \mu = i] Pr [\mu = i]$$

$$= 1 \cdot (1 - \omega) + \Pi_\mu \omega$$

$$\therefore \Pi = \omega\Pi_\mu + (1 - \omega) \qquad (3.21)$$

Thus in the absence of an advertising message the consumer would maximise expected utility using his prior beliefs Π ; making the substitution(3.20)and(3.21)into(3.10),we obtain

$$Eu = (\frac{M}{p})^{\frac{1}{2}} \{ (1 - \Pi_\mu)^2 \omega^2 p + [\Pi_\mu \omega + (1 - \omega)]^2 \}^{\frac{1}{2}} \qquad (3.22)$$

The expected value of imperfect information is given by the difference between Eu^{**} in equation(3.19)and Eu in (3.22).

$$V = (\tfrac{M}{p})^{\frac{1}{2}}\left[\{(1-\Pi_\mu)^2\ p + \Pi_\mu^{\ 2}\}^{\frac{1}{2}}\ \omega + (1-\omega) - \{(1-\Pi_\mu)^2\ \omega^2\ p + [\Pi_\mu\omega + (1-\omega)]^2\}^{\frac{1}{2}}\right] \quad (3.23$$

So $V \gtrless 0$ as

$$\{(1-\Pi_\mu)^2\ p + \Pi_\mu^{\ 2}\}\ \omega^2 + (1-\omega)^2 + 2\omega(1-\omega)\{(1-\Pi_\mu)^2\ p + \Pi_\mu^{\ 2}\}^{\frac{1}{2}}$$

$$\gtrless\ (1-\Pi_\mu)^2\ \omega^2\ p + [\Pi_\mu\omega + (1-\omega)]^2$$

i.e. as $\{(1-\Pi_\mu)^2\ p + \Pi_\mu^{\ 2}\}^{\frac{1}{2}} \gtrless \Pi_\mu$

and $(1-\Pi_\mu)^2\ p \geq 0$ as $\Pi_\mu \leq 1$

So $V \geq 0$

and the expected value of imperfect information is non-negative.

Now consider the actual value of imperfect information. We are concerned with comparing the utility received from making a decision based upon the advertising message, after the unknown variable has been realised, with the utility from making a decision on the basis of prior beliefs alone.

Suppose that $\alpha_1 = 1$ and $\alpha_2 = 0$, and further that $\mu = 1$. That is the good x_1 contains the required characteristic, and the producer of x_1 advertised this fact.

Then if the consumer had not observed the advertising message, and made the decisions in(3.9),he would consequently have realised a utility level of

$$\left[\frac{Mp\ (1-\Pi)^2}{p(1-\Pi)^2+\ \Pi^2}\right]^{\frac{1}{2}} \tag{3.24}$$

If instead the consumer had seen the advertising message $\mu = 1$, he would maximise

$$Eu = (1 - \Pi_\mu)\ x_1^{\frac{1}{2}} + \Pi_\mu\ x_2^{\frac{1}{2}}$$

$$\text{s.t.}\quad x_1 + px_2 = M$$

which yields the optimum demands

$$x_2 = \frac{M}{p} \cdot \frac{\Pi_\mu{}^2}{p(1-\Pi_\mu)^2+\ \Pi_\mu{}^2} \tag{3.25}$$

$$\text{and}\quad x_1 = M \cdot \frac{p\ (1-\Pi_\mu)^2}{p(1-\Pi_\mu)^2+\ \Pi_\mu{}^2}$$

Following this policy will yield a utility level, in the case of $\alpha_1 = 1$, of

$$\left[M \cdot \frac{p\ (1-\Pi_\mu)^2}{p(1-\Pi_\mu)^2+\ \Pi_\mu{}^2}\right]^{\frac{1}{2}} \tag{3.26}$$

Thus the actual value of imperfect information, when $\mu = 1$ and $\alpha_1 = 1$, is given by(3.26)minus(3.24).

$$V_A \Big|_{\substack{\mu = 1 \\ \alpha_1 = 1}} = M^{\frac{1}{2}} \left[\frac{p\ (1-\Pi_\mu)^2}{p(1-\Pi_\mu)^2 + \Pi_\mu{}^2} - \frac{p(1-\Pi)^2}{p(1-\Pi)^2 + \Pi^2} \right]^{\frac{1}{2}} \tag{3.27}$$

Thus (3.27) $\gtrless 0$ as $(1-\Pi_\mu) \gtrless 1 - \Pi$

$(1 - \Pi_\mu)$ is the posterior probability that $\alpha_1 = 1$, and $(1 - \Pi)$ is the prior. Thus if the posterior puts greater weight on the belief that $\alpha_1 = 1$, and in fact $\alpha_1 = 1$, then the actual value of the advertising message is positive.

Now suppose that $\alpha_1 = 0$ and $\alpha_2 = 1$. Firstly let $\mu = 1$, that is the producer of x_1 is falsely claiming that his good contains the required characteristic. Then again we can evaluate the increase in utility from the advertising message. If the consumer had made the decision in (3.9) he would realise a utility level of

$$\left[\frac{M}{p} \cdot \frac{\Pi^2}{p(1-\Pi)^2 + \Pi^2} \right]^{\frac{1}{2}} \tag{3.28}$$

If instead he had observed the advertising message and made the decisions in (3.25) he would have realised a utility level of

$$\left[\frac{M}{p} \cdot \frac{\Pi_\mu}{p(1-\Pi_\mu)^2 + \Pi_\mu{}^2} \right]^{\frac{1}{2}} \tag{3.29}$$

The expected value of imperfect information is given by (3.29) minus (3.28).

$$V_A \Big|_{\substack{\mu = 1 \\ \alpha_1 = 0}} = \left(\frac{M}{p}\right)^{\frac{1}{2}} \left[\frac{\Pi_\mu{}^2}{p(1-\Pi_\mu)^2 + \Pi_\mu{}^2} - \frac{\Pi^2}{p(1-\Pi)^2 + \Pi^2}\right]^{\frac{1}{2}} \qquad (3.30)$$

But (3.30) $\gtrless 0$ as $\Pi_\mu \gtrless \Pi$. But for (3.27) > 0 we required $1-\Pi_\mu > 1-\Pi \rightarrow \Pi_\mu < \Pi$. Thus if (3.27) > 0 then (3.30) < 0.

So if the advertising message results in the consumer adopting a new set of beliefs that have been directed by the advertising message; and if the advertising message is truthful, this will have a positive effect on actual utility. If the message is false it will have a negative effect.

This example has considered a simple consumption technology and an equally simple advertising technology. The consumer ended up purchasing one good or the other and the advertising statement was either correct or incorrect. In the more general case the consumer will purchase a number of goods, with a number of goods providing the desired characteristics in various proportions and combinations. The advertising message tells the consumer about the combination of characteristics in the advertised good. But do advertisements fulfil this role?

Advertising is no substitute for true information, or more correctly is a poor substitute for accurate information which is obtained by paying a search cost and then examining the good. Of course advertising has the advantage that it has zero cost which makes it a better substitute, but it is still an imperfect form of information.

Firms are forced to tell the truth in an advertising message because of the presence of an anti-deception mechanism; in an environment with more than a single period, the possibility of repeat purchase by the consumer, means that the firm will not find it profitable to lie; since consumers who purchase the good on the basis of a deceptive advertisement will not purchase the good in future.

The effectiveness of this mechanism depends upon a number of implicit assumptions concerning both producer and consumer behaviour. Firstly, it is assumed that the consumer is able to quantify the quality of the purchased good and compare it with some norm or its advertised value. This may be difficult; for example how does a consumer assess the quality of a good which guarantees to increase the probability of longer life. Secondly, the consumer needs to have perfect memory recall over past advertising messages, from which he constructs a black list of goods to avoid in the future. This leads on to the third assumption concerning the asymetry of demand with respect to the falsely advertised goods. For the mechanism to function, the consumer is required to positively discriminate against the advertised goods. This is a step further than simply supposing that the consumer learns to discount the information from an advertising message. The consumer develops an aversion to buying the good of the deceptive advertiser, possibly due to a chagrin effect, whereby the consumer reflects his annoyance at being duped and refuses to purchase that good in the future. This assumption implies irrational behaviour on the part of the consumer; since in the case of a high quality good whose high qualities had been exaggerated by an advertiser, the consumer would chose a good of lower quality in the future. Fourthly, the firm must have compared the present value of a truthful strategy which yields a stream of future profits, with the return from a single period; since a deceptive statement may produce high sales in the first period, but if individuals communicate about its defects will produce nothing afterwards.

Klein and Leffler (1981) consider the problem of assuring contractual performance. They note that it may be profitable for a firm to produce a low quality good and sell it at a high quality price, in terms of a once-and-for-all sale being greater than a perpetual stream of sales. However they argue that there will exist a price premium such that the profits from honest production are greater than from dishonest. This price premium is made into an equilibrium price by the existence of non-salvageable assets which have a capitalised value equal to the discounted sum of future profits. They argue that investment in brand advertising acts as a non-salvageable asset. Thus high quality is assured by the level of brand advertising.

Finally, it is assumed that the world is deterministic. If instead the world is stochastic, then verifying an advertising statement is much more difficult. Suppose the relationship between goods and characteristics is of the form

$$c = zx + \varepsilon$$

where ε is a random variable and z is the unknown parameter. Further suppose the consumer is not allowed to observe the realised value of ε directly. On purchasing an amount of the good, the consumer only observes the quantity of the characteristic produced. Thus although z has an unknown fixed value, the consumer can never be certain of its value as actual utility always depends upon the realisation of the random component. Thus an advertising statement about the value of z can never be verified.

If these assumptions do not hold, the producer is able to make a statement about his product that is untrue; though, this is perhaps a rather extreme option. It seems unlikely that the producer will make completely false claims about his product, but more realistically will be able to distort the perceived qualities of the good and present the information in a favourable light. The deviation of the actual advertising statement from the true statement will depend upon the extent of the failure of the anti-deception mechanism. For some types of goods it will be easier to make a deceptive message and escape detection. Experience goods are more likely to be accompanied by incorrect messages than search goods because of the nature of the products. The quality of experience goods can only be ascertained after purchase which allows the producer to make statements than can not be verified until after consumption has taken place. Taking account of the nature of advertising the basic premise of this chapter is that the purpose of an advertising statement is to provide information about the parameters in the consumer's decision problem. That is, the advertising message provides the consumer with a set of prior beliefs about the unknown variable. The consumer uses this information to update his initial subjective distribution. The producer makes a statement about the unknown parameter in the utility function, but the consumer does not accept this as gospel but uses the information to construct only his subjective distribution of the unknown parameter. If the consumer believes that the advertising statement is truthful then the subjective distribution is just a point estimate of the actual unknown parameter. If the consumer is more sceptical, then he believes that the actual value of the unknown parameter is only located in the vicinity of the advertiser's message.

We now state more formally how a series of advertising statements can be incorporated into a consumer's beliefs. We suppose the consumer updates his current beliefs according to Bayesian rules. His posterior beliefs are formed from his prior beliefs and any additional information.

Suppose the consumer observes a series of advertising messages $\mu_1, \ldots \mu_k$, purporting to inform him about the value of a parameter z. The consumer is assumed to know that each message is a transformation of a random variable $\tilde{\rho}_i$,

$$\tilde{\mu}_i = z + \tilde{\rho}_i \tag{3.31}$$

where $\tilde{\rho}_i \sim N(\theta, \frac{1}{\xi})$, so $\mu_i \sim N(z + \theta, \frac{1}{\xi})$. The consumer knows that each advertising message has two separable components: the true value of the unknown parameter z and, a random component which reflects both the random noise in the environment and also a deception element. Noise occurs because of problems of communication between producer and consumer. The deception element is a deliberate attempt by the producer to exaggerate the properties of his product. The consumer holds beliefs about $\tilde{\rho}_i$, which are assumed to be normal with known mean θ and degree of precision ξ. The values of these parameters are essentially endogenous, since they are concerned with the accuracy of the advertising message which is conditioned by previous attempts at deceipt. θ represents the amount, on average, which the consumer believes the advertising message will overstate product quality, and ξ reflects the strength of those beliefs. However we will suppose that both θ and ξ are given initially.

The series of advertising messages allows the consumer to infer the true value of the unknown parameter. In fact a sufficient statistic for this inference is the sample mean.

$$\bar{\mu} = \frac{1}{k} \sum_{i=1}^{k} (z + \tilde{\rho}_i) \tag{3.32}$$

If there are k advertising messages then $\bar{\mu} \sim N(z + \theta, \frac{1}{k\xi})$.

We assume that the consumer holds prior beliefs about z which are normal with mean μ_o and precision ϕ. We might expect that as these beliefs have been formed on the basis of very limited information that ϕ will be very small. The consumer constructs his posterior beliefs according to Bayes rule, yielding a normal distribution with $\hat{\mu}$ and variance $\hat{\sigma}^2$ 4 where

$$\hat{\mu} = \frac{\mu_o \phi + k\xi \bar{\mu}}{\phi + k\xi}$$

$$\text{and} \quad \hat{\sigma}^2 = \frac{1}{\phi + k\xi} \tag{3.33}$$

These updated beliefs will be used as the parameters in the consumer's decision problem in (3.4).

The mean and variance stated in equation(3.33)result from the assumptions of normality with regard to the likelihood function and the prior. In a more general case, these two functions could be described by any distribution, whichever best captures the characteristics of ignorance with respect to the prior, and the deception element with respect to the series of advertising messages. Though in defence of the distributions chosen here it can be seen that even if the series of advertising messages gives the same value, the variance of the posterior in (33) gets reduced the more messages the consumer observes.

We can now distinguish between two types of information: unbiased information which the consumer obtains from searching, and information from advertising which may be distorted. The role of advertising as an information generating technique can be incorporated into the consumer's decision problem in the following way. It is postulated that the consumer faces a two stage sequential decision process. Initially the consumer is in ignorance and holds diffuse beliefs about an unknown parameter in the utility function. In the first stage, information is provided by advertising, which allows the consumer to update his diffuse distribution. The consumer now holds new beliefs about the unknown variable which are concentrated around a point as guided by the advertising message which may or may not be correct. The consumer can choose to make an allocation decision on the basis of this incomplete information set or can choose to move into the second stage and obtain further information through searching. In the case of experience goods, we would expect the consumer to make a decision to purchase or not at the end of the first stage. In the case of search goods, where the costs of search are low the consumer would probably move into the second stage before deciding upon the quantity to buy.

If the utility function is stochastic, the consumer has another type of information available as an alternative to search, and that is information due to sampling. When the consumer purchases a quantity of a good, he is taking a random sample, and gains information about the underlying distribution from the sample chosen. Grossman, Kihlstrom and Mirman (1977) show that a consumer who is aware of this information due to sampling, will choose a larger sample than otherwise because of the additional information. The theory of search when price is unknown has been well documented. The original treatment was given by Stigler (1961). He considered the optimal number of observations an individual should make in trying to find the lowest price from a variety of stores. The individual was assumed to have a probability distribution over the possible range of prices. This strategy has been termed fixed-sample-size strategy (by Manning and Morgan, 1982) and when observations are observer sequentially it has been shown by McCall and DeGroot to be an inferior strategy to sequential search. (by McCall (1970) and DeGroot (1970)) Under sequential search the individual holds a reservation price, and continues to search until he finds a price less than or equal to the reservation price. Kohn and Shavell (1974) distinguish between adaptive and static search. Under static search, the individual holds rigid beliefs so that his reservation price remains constant over the search period. Under adaptive search the consumer updates his beliefs on the basis of past observations and his reservation price changes in the light of his experience.

The problem with these papers from the viewpoint of quality uncertainty, is that they consider a one-off purchasing decision. The consumer searches for the lowest price and then purchases the good, with the implicit assumption that any further purchases will take place at that price. If product quality is a random variable, then having observed one unit of the good does not guarantee subsequent knowledge. In the next period the consumer will be faced with a new problem, except that he knows what gave him the highest level of quality last time. Under these circumstances sampling may prove a more effective, and certainly cheaper form of search strategy.

Consider the problem facing the consumer at the start of the second stage. He has an unknown parameter z in his utility function: $U(x, z)$. The subjective distribution function is $F(z)$ and density function $f(z)$. This density function is normal with parameters $\hat{\mu}$ and $\hat{\sigma}^2$, given by (3.33). The consumer's decision problem is to choose an x which maximises expected utility

$$\max_{x} \quad Eu = \int_{z} U(x, z)\, f(z\,|\,\hat{\mu}, \hat{\sigma}^2)\, dz \tag{3.34}$$

Again, the consumer could compute the expected value of (perfect) information, by working out the expected utility of a fully informed decision, and compare with the costs of information. We could perhaps guess that the expected value of information will be less as the consumers degree of confidence gets larger. Though, to repeat an earlier point made by Gould, the maximum value of information does not necessarily involve equiprobable outcomes. The implication being that advertising statements which concentrate beliefs about the unknown parameter around a point increase the expected value of perfect information, which makes it more likely that a consumer will engage in information search.

How does this affect the distinction between inspection and experience goods? For search goods the costs of inspection are low so even if the expected value of information is low, the consumer may still engage in search. Any false claims by an advertising statement could be quickly invalidated. The advertiser is aware of this possibility so we may expect for inspection goods that advertising statements reflect the true values of the unknown parameters. For experience goods the costs of search are high and skilled advertisers may be able to affect the subjective distribution function in such a way as to minimise the expected value of information so that the consumer is less likely to engage in active search. It should be noted that Gould's conclusions do not imply that equiprobable outcomes minimise the expected value of information nor single point estimates maximise it. If the consumer is less likely to engage in search then the producer has more degrees of freedom to twist the information in his favour.

5. The Informative Content of Advertising Messages

The literature has been concerned with the distinction between informative advertising and persuasive advertising. The importance of this distinction is in its effect on the elasticities of demand. If advertising is informative (Nelson, (1974), and Ferguson, (1974)), then advertising messages tell consumers about alternative products that are available, elasticity of demand is a function of known substitutes so that the more substitutes that the consumer is made aware of, the more elastic is his demand curve for a particular product.

Persuasive advertising (Galbraith, (1967)) attempts to mould consumer preferences. If successful it will create brand loyalties which will reduce demand elasticities. Brand loyalty can be defined as an inelastic demand for a particular product. Ferguson claims "Advertising makes demand curves more elastic by providing consumers with information about brand qualities".[5] Although the demand curve for a good will be more elastic the more substitutes are available, or the more substitutes that are known to be available, the demand curve for the good that is being advertised will become more inelastic. Even if advertising is providing information about quality uncertainty, by affecting unknown parameters in the consumer's utility function, Basman (1956) shows that the marginal utility of the good will be increased as these parameters vary, which is regarded as persuasive advertising. An additional problem of stating that the purpose of advertising is solely to inform consumers of available substitutes in the same industry, is that advertising is assumed to have no effect on the industry demand curve. The industry demand curve is determined by the locus of equilibrium positions for the consumers between changes in prices of the industry product relative to the products of other industries. The fact that a new brand of product has become available will only alter intra-industry demand. The share of the industry demand curve facing each individual firm may be affected; specifically after the introduction of a new product it will shift backwards which ceterus parabus will make it more elastic. This increase in elasticity is due to the same industry output being divided between more firms. Consequently this type of information does not explain why monopolists advertise. If advertising only provides information about competing brands a monopolist has no incentive to advertise.

Dixit and Norman (1978) impose a parameter on the utility function, and assume it is an increasing function of the level of advertising. They analyse the effects of changes in the parameter on welfare. By supposing that it is the level of advertising which determines the value of the parameter, Dixit and Norman consider the effect of a small change in the level of advertising on the utility function. The result is a small shift in the demand curve. As the shift in the demand curve is only slight, the increase in consumer surplus can be ignored, which means there is no net welfare increase from advertising. However, if the value of the parameter depends upon the information contained in the advertising statement and not just upon the quantity of advertising, then a small change in the level of advertising may bring about a large shift in the demand curve, and in this case the increase in consumer surplus cannot be ignored and will represent the increase in welfare from an advertising message.

They provide no rationale for the inclusion of the parameter affected by advertising. According to Kotowitz and Mathewson (1979) "specifying advertising in the utility function is not an acceptable way to proceed. It offers no understanding of the role of advertising in the consumer's decision process. If all the characteristics of a commodity were readily measurable and could be validated, then they would be known to the consumer. In this case why should advertising increase the consumer's marginal utility for the good?"[6] One view of advertising is that the content of an advertising message is not so important as the quantity of message. Nelson (1974) originally, and more recently Klein and Leffler (1981) argue that advertising only provides indirect information.

Nelson (1974) extends ideas developed in an earlier paper (Nelson, 1970). The distinction between search and experience goods is a necessary prerequisite for his subsequent conclusions. Given that consumers are ignorant of a product's qualities, advertising provides information, but the nature of the information depends upon whether the product is a search or experience good. Advertising provides direct information for search goods but only indirect information for experience goods. What is the distinction between direct and indirect information and how do they relate to search and experience goods?

Nelson explains that because search goods can be examined prior to purchase, the purpose of advertising is to entice consumers to sample the product. In examining the good they will be able to verify the advertising statement. If the firm has made exaggerated claims about its product quality the consumer can choose not to buy the product and will remember to ignore any future advertisements from this particular firm. The firm loses its credibility, and the loss of credibility acts as an anti-deception mechanism in ensuring that no false claims are made about product quality. Therefore advertising of search goods will provide direct, accurate information. Nelson contrasts this situation with experience goods where as advertising statement can not be verified prior to purchase. In this case "Advertising provides no direct information ... (By direct information I mean information contained in the advertising statement.) After Pepto Bismol has been correctly identified as a stomach remedy, the statement that Pepto Bismol is most soothing is without information content. Its producers have an incentive to say so even if it were the least soothing of stomach remedies." [7]

But although advertising statements of experience goods do not have any information content, they are a form of indirect information since consumers learn that the more a firm advertises the higher the quality of the product. The reason why more advertising implies a better buy is demonstrated by a number of scenarios. The argument is basically that firms with a high quality product have an incentive to advertise. Again the anti-deception mechanism is the fact that consumers will at some stage in the future repeat their purchases. This is what Nelson terms "consumer power in the product market".[8] Thus Nelson distinguishes between two types of advertising, advertising of search goods which contains direct information, and advertising of experience goods which is indirect and contains little "hard" information, and whose function is mainly to enhance the reputability of a brand.

Similarly Klein and Leffler show that consumers recognise that the quantity of advertising messages are a non-salvageable investment by the firm. Firms who cheat, that is, produce low quality products and claim they are of high quality, sacrifice this sunken investment. Thus large amounts of advertising are guarantees of high quality. However, even if large advertising expenditures are a sufficient condition for not cheating, they are unlikely to be a necessary condition since some small honest firms may be unable to raise the finance to engage in expensive advertising campaigns. Of course in this case the situation may degenerate into all small firms producing low quality goods, which is analagous to Akerlof's market for lemons (1970). Even if large levels of brand advertising are necessary and sufficient conditions for "honest" production, how do consumers judge between two rival advertisers? The knowledge of a "fair" price is not sufficient for the consumer to maximise his expected utility. To some extent the consumer must observe the direct information inherent in an advertising statement.

Similarly the objection to Nelson's analysis is that advertising statements of experience goods have little information content. Is is one thing to argue, as the above quote suggests, that consumers realise that producers have an incentive to make wild claims about their products, but that does not imply that consumers consequently discount advertising messages entirely. In the absence of any other information, these messages will be the only facts that consumers can base their decisions on, and to argue that they will have no effect on consumer behaviour is unconvincing. Nelson himself in one of his tests found "six brands claiming that they were the largest brands for their respective product classes".[9] But this is a statement with information content: the brands are saying that in being the largest brands, "surely so many people who buy our product cannot be wrong?" Of course this is not a particularly hard piece of information in the sense that it informs consumers how good the brand is, but that, as already noted, is not the advertisers' concern. The firm is concerned with selling its product. It may be aware that it cannot make false statements about product quality because of what Nelson terms "consumer power". In this case the firm can try to wriggle around the problem by giving information that is not quantifiable and is difficult to ascertain. The advertisers art is to make statements that will improve the image of the product but are sufficiently elusive to remove the possibility of any comeback. The information of these messages will be presented in a highly persuasive way.

Some of Nelson's findings can be interpreted in different ways. For example, Nelson finds that producers of experience goods advertise more than producers of search goods which supports his hypothesis that advertising of experience goods increases sales through increasing the reputation of the seller. However, this same evidence could be interpreted as showing that experience goods advertising needs to be more intense than advertising for search goods, since perhaps consumers are more sceptical of them.

Similarly Nelson finds a higher concentration of search goods advertising in newspapers as evidence that the direct information content of search goods advertising means that consumers may need to refer back to the information source. However the same evidence would satisfy the hypothesis that firms advertising experience goods which may involve an element of deceit, need to "flash" their advertisements on television screens, so that consumers have difficulty apportioning the blame for purchasing a good which was not up to the standard they had believed.

6. Conclusions

In this chapter we have attempted to place advertising in the role of an information generating mechanism. Imperfect information is represented by supposing that one or more parameters in the consumer's function are unknown. Advertising provides the consumer with information to enable him to construct a subjective distribution function to describe the possible values of the unknown parameter. Thus the advertising message has value, and the value is equal to the increase in consumer surplus that results from the consumer making a fully informed decision as opposed to a decision make in ignorance, which will probably mean that the consumer is not equating the actual marginal evaluation of the good with its price.

However, advertising may not give accurate information since it is provided by the person selling the good. Advertising as information can be contrasted with the unbiased information which is generated by random sampling or search. The failure of repeat purchase as an anti-deception mechanism due to imperfections in the environment, means that advertising messages may contain an element of deception. Thus advertising may or may not have a positive value, but its value can again be calculated by evaluating the utility that arises from a random sample and the utility from the consumption of a good based on an advertising message.

Whatever the actual value of advertising, the consumer will base his decision to seek further information or not, on the expected value of information. We have argued that advertising results in the consumer moving from an initial diffuse subjective distribution describing the unknown parameter to a modal one. If the advertising message is untruthful or exaggerated, it will be in the producer's interest to present the information in such a way as to minimise the expected value of further information, so that the consumer will make a decision solely on the basis of the advertising statement.

How will advertising affect consumer demand? We wish to stress the distinction between the quantitative and qualitative aspects of advertising. The former refers to the amount of advertising or the number of advertising messages, the latter to the nature of the advertising statement. When we speak of an increase in advertising it is normally taken to mean an increase in the number of advertising messages. But while it is likely that an increase in the number of messages will increase either the proportion of the population who see the statement, or the probability that any one individual will see it; from the point of view of an individual who has already viewed the message, more of the same is unlikely to have any effect. Though to repeat an earlier point, repetition of the same statement may affect the consumer's beliefs about the content of the message as given by equation (3.34).

However, if the change in advertising refers to a change in the advertising statement, then we would expect this to have an effect on consumer demand. If advertising affects the subjective value of a parameter in the utility function, and if this increases the marginal utility of the good then the demand curve must become more inelastic. It does not matter whether the advertising statement is "informative" or "persuasive", in either case the demand curve for that good will have a lower elasticity. Of course, the effect of this statement on the differentiated products of other producers may make their demand curves more elastic. Conversely, we would expect that other unknown parameters in the utility function of the consumer will be affected by rivals' advertising messages which will make the demand for the good in question more elastic.

Notes

1. Mood, Graybill and Boes (1974)
2. This proposition and proof was pointed out to me by Jane Black.
3. Telser (1966), P.460.
4. For derivation see DeGroot P.167.
5. Ferguson (1974), p.31.
6. Kotowitz and Mathewson (1979), p.567.
7. Nelson (1974), p. 731.
8. p. 730.
9. p. 733.

Chapter 4

Learning

1. Introduction

This chapter consider the importance of a consumer's ability to learn in an environment with imperfect information. It illustrates how in a two period allocation problem with uncertainty in each period, the consumer's decisions are influenced by the knowledge that he is able to learn about the uncertainty. The time periods are linked through the learning process of the consumer.

The problem to be analysed is that faced by a consumer deciding whether or not to purchase a new good whose quality is unknown. The consumer's initial beliefs about product quality have been provided by a series of advertising statements. Because of the two period environment, the consumer is able to experiment with the new good in the first period, and verify the advertising messages, before making a final decision in the second period. Given this opportunity for learning, how will the consumer's decisions be affected in the first period?

We can distinguish two types of learning: the acquisition of knowledge and the acquisition of skills. The former implies that the consumer initially has imperfect knowledge, whereas the second definition implies some kind of investment procedure, whereby learning is a current imput into the stock of ability. This second definition has been coined "learning-by-doing", and was suggested by Arrow (1962). He imposed the assumption of a learning time trend on an economic growth model such that investment in the latest technology yields greater productivity, since the latest technology reflects the cumulated skills of previous technologies. This type of learning is also suggested as a source of economies of scale.

Here, we are concerned with the first definition of learning, which could be called learning-by-sampling rather than learning-by-doing. The effect of introducing a learning mechanism is that the consumer is able to accumulate knowledge about the quality of the new product, and in the long run we can envisage this learning as complete when the consumer has acquired perfect information. Thus the importance of introducing learning into a model of consumer behaviour is in its short run effects as the consumer adjusts towards his long run equilibrium position.

In line with this view of learning, we firstly consider the optimal decisions for the consumer in a two period environment. In the final period it is assumed that the consumer has acquired all the information it is ever likely to, but in the first period, representing the short run, the consumer has less than complete information and must decide whether to gain information or pursue single period utility maximisation.

By assuming that there is an element of uncertainty in the environment, which can be represented by a distribution function, we can incorporate information into the model by supposing that a parameter in the environment is unknown. In a deterministic world, any imperfect knowledge could be overcome by observing the unknown parameter. This is of course an interesting problem and the search theory literature is concerned with answering it. However, our assumption of a stochastic world means that an observation does not necessarily imply knowledge of the unknown parameter. For example, an observation on the quality of an umbrella depends upon the severity of the weather, which is a random variable: observing the response of the umbrella to a bout of mild weather does not give complete information of the quality of that umbrella. Defining learning as the acquisition of information with respect to the parameters in the distribution function, means that the learning process can be modelled using the statistical techniques of Bayesian analysis.

Kohn and Shavell (1974) consider the problem of sequential decision making. They define the problem when the distribution function has known parameters as static. There is no possibility for learning in the static case. Any other type of sequential decision making is adaptive. Bayesian learning about an unknown parameter of the distribution function is an example of the adaptive case. We shall be concerned with examining the consumer's decision problem under both static and adaptive distributions.

In two illuminating articles in 1974, Kihlstrom uses Bayes' rule to solve the consumer problem of maximising utility in a single period when product quality is a random variable with an unknown distribution. The consumer is allowed to purchase a sample of the good, and in experiencing the sample gains information about product quality which enables him to update his subjective distribution function.

Grossman, Kihlstrom and Mirman (1977), provide a more general approach to adaptive sequential decision problems. They explicitly recognise the two period nature of the problem: sampling occurs in the first period, before a final decision is made in the second. The solution is found by dynamic programming, using the technique of backward induction. The consumer maximises utility in the second and final period for every possible realisation of the random variable from the first period. These optimal values are then weighted by the probability that the random variable will take on any particular value. The consumer then maximises the discounted sum of expected utilities in the first period, which includes an indirect utility function for the second period.

This chapter attempts to use the general framework proposed by Grossman et al. and apply it to a linear utility function. We find that the Grossman et al. results hold, provided the budget constraint is exhausted in each period. When we consider a flexible budget constraint there is a possibility that the earlier results will be reversed. We look at various comparative static exercises and also examine the implications of sampling on the efficiency of the household technology when the consumer is learning about a new good which has an associated demand for a complementary product.

2. The general model

A consumer gains utility in each time period from two characteristics c_t and c_{ot}. The former is produced by consuming a good x, but also depends upon the realisation of a random variable, ε and a parameter z, which is fixed through time, but whose value is unknown by the consumer.

$$c_t = c_t\,(x_t, z, \tilde{\varepsilon}_t) \tag{4.1}$$

The other characteristic is produced directly by consuming a good y : $c_{ot} = y_t$. The consumer has a strictly concave utility function which is invariant over time. However the utility in any period will depend upon the characteristics consumed in that period:

$$u_t = u(y_t, c_t(x_t, z, \tilde{\varepsilon}_t)) \tag{4.2}$$

The consumer faces a budget constraint in each period

$$M = y_t + px_t \tag{4.3}$$

where p is the relative price of good x , since the price of y has been normalised to 1.

$\tilde{\varepsilon}_t$ is a sequence of independently and identically distributed random variables with a density function f_{ε_t} . The consumer holds beliefs about the possible values of z represented by a subjective density function f_z^t .

The time superscript shows that f_z changes over time, whereas it is assumed that the distribution of the $\tilde{\varepsilon}_t$ remains fixed over time.

When the consumer purchases a quantity of the goods x_t and y_t , he will be subjected to the realisation of the random variable and observes the subsequent quantity of the characteristic c .

The consumer is not allowed to observe the random variable directly since this would enable him to compute z immediately. The consumer's experience is summarised by a vector:

$$\underset{\sim}{w}_t = \underset{\sim}{w}_t(c_t, x_t) \tag{4.4}$$

Observing this vector of statistics, which are defined as being sufficient[1] in the sense that they contain sufficient information to make an inference about the true value of z, allows the consumer to update his initial beliefs on the basis of his experience. If the consumer follows Bayesian rules, then he updates the subjective density function of the unknown parameter according to:

$$f_z^t = f_z^t(f_z^{t-1}, \underset{\sim}{w}_{t-1}) \tag{4.5}$$

That is, current beliefs depend upon last periods beliefs and the experience from the last period. But from (4.4), the experience in $t-1$ depends upon x_{t-1}. Thus the time periods are linked through the learning process of the consumer. Not only does the purchase of x_t contribute directly to current utility, through the quantity of characteristic c, but the purchase of x_t provides information to the consumer which he uses to update his subjective density function in period $t + 1$.

A consumer making a decision in period t, recognises that the decision will affect utility in the next period. The consumer therefore makes a decision in the current period to maximise the expected value of the discounted sum of utility over two periods

$$\max \ EU_1 = \sum_{t=1}^{2} \frac{1}{(1+i)^{t-1}} \ Eu_t \qquad (4.6)$$

This problem can be solved by dynamic programming. The consumer solves his optimal plan over two periods by considering the problem in the final period first of all, and then working backwards.

Consider the problem in the final period. Since period 2 is the final period, there is no future and the decision that the consumer makes will only affect utility in the final period; it will have no further effects. Thus the consumer maximises expected utility as a single period problem, conditional upon the previous value of x, and the realisation of the statistic $\underset{\sim}{w}$ in the previous period. Let V_2 be the maximum value function in the final period, then

$$V_2 = \max_{y_2, x_2} \int_{\forall z} \int_{\forall \varepsilon_2} u(y_2, c_2(x_2, z, \tilde{\varepsilon}_2)) \, f_{\varepsilon_2} \, f_z^2 \, d\varepsilon_2 \, dz \qquad (4.7)$$

$$\text{s.t.} \quad M = y_2 + px_2$$

where, from (4.4) $f_z^2 = f_z^2(f_z^1, w_1(x_1, c_1(\varepsilon_1, z, x_1)))$ which is termed the posterior distribution of z. But looking from the first period ε_1, has yet to be realised, so $\underset{\sim}{w}_1$ are a vector of random variables with joint density function $f^1_{w_1}$. Thus the consumer will carry out the maximisation procedure implied in (4.7) for every possible realisation of the random variable $\underset{\sim}{w}_1$. Having obtained the optimal decision variables he will weight the subsequent indirect utility by the probability that any particular values of the $\underset{\sim}{w}_1$'s will obtain:

$$EV_2 = \iint_{\forall \underset{\sim}{w}_1 \cdots\cdots} \cdots\cdots V_2 \; f^1_{w_1} \, d\underset{\sim}{w}_1 \qquad (4.8)$$

The final period expected indirect utility function can be written as $EV_2\ (x_1)$, showing its dependence on the decision variable in the previous period.

The consumer having solved the final period problem moves back one period and solves the following problem in the first period.

$$\max_{x_1, y_1} \quad EU_1 = Eu_1 + \frac{1}{1+i} EV_2(x_1) \qquad (4.9)$$
$$\text{s.t. } M = y_1 + px_1$$

where i is the social rate of discount, and Eu_1 is given by:

$$Eu_1 = \int_{\forall z} \int_{\forall \varepsilon_1} u(y_1, c_1(x_1, z, \tilde{\varepsilon}_1)) \, f_{\varepsilon_1} f^1_z \, d\varepsilon_1 \, dz$$

In the first period, the consumer uses his prior distribution of z, f^1_z, as the expectations operator over the unknown parameter. We argued in Chapter 3 that this prior distribution is constructed from information given in advertising messages

> "Analysis of the advertising process begins by assuming that consumers have initially a stock of knowledge about goods and services of which part comes via advertising messages."
>
> Telser (1966, p.462)

Advertising furnishes the consumer with the parameters in the prior distribution of the unknown variable. It should be noted that this prior distribution is still present, after the learning process has occurred, in equation (4.7), though we shall see that its weighting has changed.

3. The Model with a Linear Utility Function

The utility function is written as:

$$u_t = y_t + c_t \tag{4.10}$$

The characteristic c is produced by consuming x units of a new good, but the quality of each unit is a random variable $\tilde{\varepsilon}_i$:

$$c_t = \sum_{i=1}^{x_t} \tilde{\varepsilon}_{it} \tag{4.11}$$

For example if the consumer purchases 1-lb of apples then $\tilde{\varepsilon}_{it}$ represents the quality of each apple. If $-\infty \leq \tilde{\varepsilon}_{it} \leq +\infty$, where $\tilde{\varepsilon}_{it}$ are a sequence of independent and identically distributed random variables, then we may write

$$Eu_t = y_t + E(\varepsilon).x_t \tag{4.12}$$

The other good y_t can be thought of as either a good which is not subject to random elements or that it has a similar structure to x but the expected value of the random variable in this case is unity.

The problem for the consumer can be stated in a two period dynamic programming framework. The consumer maximises EU_1, where

$$EU_1 = y_1 + E(\varepsilon).x_1 + \frac{1}{1+i} . EV_2 \tag{4.13}$$

$$\text{s.t.} \quad M_1 = y_1 + px_1$$

and $$EV_2 = \max_{x_2, y_2} \{ y_2 + E(\varepsilon).x_2 \}$$

The distribution of each $\tilde{\varepsilon}_i$ is normal with mean z and unit variance, but the value of $z > 0$ is unknown to the consumer, since x is a new good which has not been tested before. The consumer holds subjective beliefs about the distribution of z, initially these beliefs are described by a prior normal distribution with mean μ_o and degree of precision ϕ. The degree of precision is the inverse of the variance, it reflects the consumer's confidence in the value of μ_o being the true mean: the more confident is the consumer that μ_o is the true mean, so the higher is ϕ, and the higher is ϕ the lower is the variance.

We suppose that the value of μ_o has been given to the consumer by an advertising message. In this case ϕ reflects the consumer's degree of confidence in the accuracy of the advertising message.

We argued in Chapter 3 that the distribution used in the subsequent decision problem would be given by a weighted average of the consumer's initial diffuse beliefs and his experience of a series of advertising messages. Thus the posterior from this updating procedure becomes the prior in the problem considered in this chapter. Further note, that by specifying the advertising message as μ_o, we are ignoring the previous updating process from diffuse beliefs and advertising, and in effect claiming that advertising is very heavily weighted in the prior distribution. In fact this is perhaps not an unrealistic assumption, since in the case of genuine initial ignorance all information can only come through advertising statements.

In the final period the subjective beliefs concerning z are described by posterior normal distribution. This posterior distribution is constructed from the prior and the consumer's experience in the first period.

The consumer purchases a quantity of the new good x_1 in the first period, and consequently observes the realisation of a sequence of random variables. He then forms a likelihood function from this sample of realised values; that is, the consumer asks himself from which particular distribution is it likely that this observed sample originated. Armed with this experience the consumer modifies his initial beliefs using Bayes' rule:-

Posterior distribution $\propto$ Prior distribution x Likelihood Function

The consumer obtains an updated set of subjective beliefs for use in the second period, represented by the posterior distribution.

In fact the consumer may not observe the sample of realised values, but observe the average level of product quality $\bar{\varepsilon}$, where

$$\bar{\varepsilon} = \frac{1}{x_1} \sum_{j=1}^{x_1} \tilde{\varepsilon}_j \tag{4.14}$$

and $\tilde{\varepsilon}_j$ as above has mean z and unit variance; so $\bar{\varepsilon}$ will also be a normal random variable with mean z and variance $\frac{1}{x}$, and will lie in the range $(-\infty, +\infty)$. $\tilde{\varepsilon}$ will be a "sufficient statistic" to provide information on the true mean of ε, and can be used instead of the likelihood function to compute the posterior distribution of z.

The posterior will also be a normal distribution with mean and variance given below[3]

$$\mu_{\bar{\varepsilon}} = \frac{\mu_0 \phi + \bar{\varepsilon} x_1}{\phi + x_1} \tag{4.15}$$

$$\sigma^2_{\bar{\varepsilon}} = \frac{1}{\phi + x_1} \tag{4.16}$$

The mean of the posterior distribution, is simply a weighted average of the initial beliefs μ_o, provided by the advertising message and the sample mean $\bar{\varepsilon}$.

In the first period, in solving equation (4.13) the consumer will use the expectations operator based on the prior distribution over the unknown parameter, $f(z\,|\,\mu_o, \frac{1}{\phi})$. In the final period he will use the posterior distribution $f(z\,|\,\mu_{\bar{\varepsilon}}, \sigma^2_{\bar{\varepsilon}})$. But looking from the first period before a decision has yet been made, the random variables have not yet been realised, and the sample is still unobserved; so $\bar{\varepsilon}$ is itself a random variable which will need to be integrated out. The marginal distribution of $\bar{\varepsilon}$ is also normal, $f(\bar{\varepsilon}\,|\,\mu_o, \frac{1}{\phi} + \frac{1}{x_1})$. [4]

The consumer now solves the problem in (4.13), where the posterior expectation of z is given by (4.15).

4. The solution to the model

In the final period, there is no future, so there is no utility to be gained from further information. The consumer with a linear utility function will purchase either the new good or the old good. The consumer maximises expected utility in the second period for all values of $\bar{\varepsilon}$, as in equation (4.7). Making the substitution implied by the budget constraint

$$y_2 = M_2 - px_2$$

$$\max_{x_2} Eu_2 = M - px_2 + E(\varepsilon)\, x_2 \tag{4.17}$$

where $E(\varepsilon) = z$ and $E(z) = \mu_{\bar{\varepsilon}}$ F.O.C. yield

$$\frac{dEu_2}{dx_2} = -p + \mu_{\bar{\varepsilon}} \tag{4.18}$$

$$\left.\begin{aligned} \text{So} \quad x_2 &= \frac{M_2}{p} \quad \text{if} \quad p \le \frac{\mu_o\phi + \bar{\varepsilon}x_1}{\phi + x_1} \\ x_2 &= 0 \quad \text{if} \quad p > \mu_{\bar{\varepsilon}} \end{aligned}\right\} \tag{4.19}$$

The optimal value of x_2 depends upon the realised value of $\bar{\varepsilon}$. For sufficiently high values of $\bar{\varepsilon}$: $x_2 = \frac{M_2}{p}$; for low values of $\bar{\varepsilon}$: $x_2 = 0$. The level of utility in the second period depends upon the choice of the two goods. Substituting for the optimal values of the decision variables, we obtain the indirect utility function V_2.

If $p \le \mu_{\bar{\varepsilon}} \rightarrow x_2 = \frac{M_2}{p} \rightarrow V_2 = \mu_{\bar{\varepsilon}} \cdot \frac{M_2}{p}$

If $p > \mu_{\bar{\varepsilon}} \rightarrow x_2 = 0 \rightarrow V_2 = M_2$

These optimal values of x_2, y_2 depend upon $\bar{\varepsilon}$, through equation (4.15), which is itself a random variable when viewed from the first period, so the consumer computes a maximum value function for every possible value of $\bar{\varepsilon}$, and then weights the resulting indirect utility functions by the probability of observing a particular $\bar{\varepsilon}$.

$$EV_2 = \int_{\bar{\varepsilon}^*}^{\infty} \frac{M_2}{p}\left[\frac{\mu_o\phi + \bar{\varepsilon}x_1}{\phi + x_1}\right] f\left(\bar{\varepsilon}\,\middle|\,\mu_o, \frac{1}{\phi} + \frac{1}{x_1}\right) d\bar{\varepsilon} + \int_{-\infty}^{\bar{\varepsilon}^*} M_2\, f(\bar{\varepsilon}|\cdot)\, d\bar{\varepsilon} \tag{4.20}$$

where $\bar{\varepsilon}^*$ satisfies $p = \dfrac{\mu_o\phi + \bar{\varepsilon}^* x_1}{\phi + x_1}$

Equation (4.20) can be substituted into equation (4.13) to obtain the first period objective function in which the only decision variable is x_1, and the effect of x_1 on the second period expected indirect utility function is explicitly recognised.

$$\text{Consumer} \quad \max_{x_1} EU_1 = M_1 - px_1 + E(\varepsilon)x_1 + \frac{1}{1+i} EV_2 \qquad (4.21)$$

where $E(\varepsilon) = z$ and $E(z) = \mu_o$ [5]

First order conditions yield:

$$\frac{dEU_1}{dx_1} = -p + \mu_o + \theta = 0 \qquad (4.22)$$

$$\text{where} \quad \theta = \frac{1}{1+i} \cdot \frac{dEV_2}{dx_1}$$

We wish to show that the opportunity for learning will cause the consumer to purchase more of the new good in the first period than otherwise. Firstly it is necessary to show that EV_2 is increasing in x_1 .

<u>Proposition 4.1</u> EV_2 is a non-decreasing function of x_1 .

<u>Proof:</u> The distribution of the random variable in equation (4.20) can be transformed into a standard normal distribution.

Define a new random variable $\tilde{m}$: $-\infty \leq m \leq +\infty$

$$m = \frac{\bar{\varepsilon} - \mu_o}{\sqrt{\frac{1}{\phi} + \frac{1}{x_1}}}$$

$$\text{Then} \quad \frac{\mu_o\phi + \bar{\varepsilon}x_1}{\phi + x_1} = \mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi + x_1)^{\frac{1}{2}}}$$

$$\text{and} \quad m^* = (p - \mu_o)\,(\phi + x_1)^{\frac{1}{2}}\,\left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \qquad (4.23)$$

Substituting these values in equation (4.20):

$$EV_2 = \frac{M_2}{p} \int_{m^*}^{\infty} \left[\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi + x_1)^{\frac{1}{2}}}\right] f(m\,|0,1)dm + M_2 \int_{-\infty}^{m^*} f(m\,|\cdot)dm \qquad (4.24)$$

m^* contains x_1 as an argument, but using Leibniz' rule for differentiating around an integral by evaluating the integral at m^*:

$$- M_2 \, f(m^*) \frac{dm^*}{dx_1} + M_2 \, f(m^*) \frac{dm^*}{dx_1} = 0$$

$$\therefore \quad \frac{dEV_2}{dx_1} = \frac{M_2}{2p} \cdot \frac{1}{(\phi + x_1)^{3/2}} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \int_{m^*}^{\infty} m f(m) dm > 0 \qquad (4.25)$$

Q.E.D.

Purchasing the new good in the first period increases expected utility in the second. This comes about because the larger is x_1, the larger is the sample upon which the sample mean is based. The larger is x_1 the more confident is the consumer that the sample mean is the true mean.

We can now state Theorem 1 which is the initial result of this chapter, and compares the size of the optimal purchase of the new good in the first period when the distribution function is adaptive, and learning is allowed, with the optimal purchase under a static distribution when there is no opportunity for learning.

Theorem 4.1 If the objective function is given by equation (4.21), and if x_1^o is the value of the decision variable which maximises the objective in the non-adaptive case and x_1^* is the optimal decision rule in the adaptive case, then $x_1^* \geq x_1^o$.

This result is the same as that obtained by Grossman et al., (Theorem 2, p.538).

Proof: Under a non-adaptive distribution, the time periods are independent. The consumer does not update his distribution function, the initial density function is used by the consumer in all subsequent periods.

In both the adaptive and non-adaptive cases, the consumer starts with the same prior density function. The non-adaptive case reduces to a series of single period optimisation problems:

In a single period problem, F.O.C.'s for maximum utility are:

$$\frac{dEu_1}{dx_1} = -p + \mu_o \qquad (4.26)$$

$$\left.\begin{aligned} &\text{If } p < \mu_o \rightarrow x_1 = \frac{M_1}{p} \\ &\text{If } p \geq \mu_o \rightarrow x_1 = 0 \end{aligned}\right\} \qquad (4.27)$$

The form of the utility function means that in a single period problem the two goods are perfect substitutes for each other. A set of indifference curves will be a series of downward sloping straight lines. The consumer spends his entire income on either the new good or the old good, depending on the quality relative to the price.

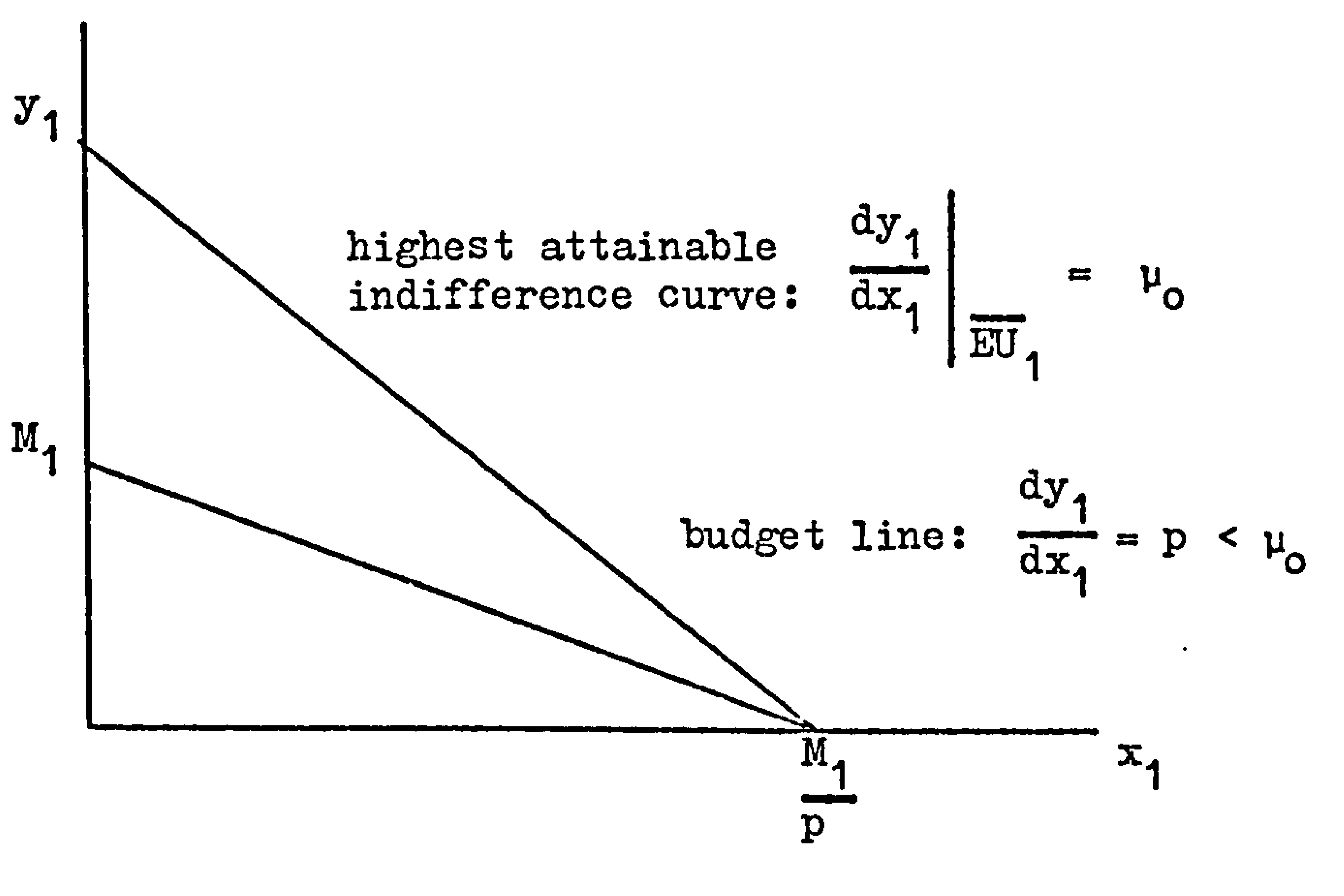


Diagram 4.1

In the adaptive case, there is an additional positive term θ, c.f. (4.22) with (4.25). There are three conditional demand schemes:

(i) If $p \le \mu_o \rightarrow x_1^* = \frac{M_1}{p}$

and $x_1^o = \frac{M_1}{p}$

(ii) If $p \ge \mu_o$ but $p - \mu_o \le \theta \rightarrow x_1^* \le \frac{M_1}{p}$ with complementary slackness

but $x_1^o = 0$

(iii) If $p > \mu_o$ but $p - \mu_o > \theta \rightarrow x_1^* = 0$

$x_1^o = 0$

Thus $x_1^* \ge x_1^o$, and it is condition (ii) that ensures the inequality sign.

Q.E.D.

In a single period model, or if the consumer's outlook was myopic and did not recognise its ability to learn about its uncertain environment, if the product price was greater than the expected value of the product quality, the consumer would purchase none of the new good. The additional positive term on the right-hand side of equation (4.22) means that it is more likely that the consumer will purchase the new good in the adaptive case.

The reason that $x_1^* \ge x_1^o$ is that in the adaptive case not only does the consumer obtain utility directly from consuming the good, but also gains information about the quality of the new good, information which will be used in the next period. Information and the new product are complementary goods, if the consumer ignores this complementarity his expected utility will be lower.

From equation (4.24) it can be seen that EV_2 depends upon x_1, so that there is the possibility of an internal solution. It may well prove optimal for the firm to mix the two goods in the first period, whereas in a myopic model, this would never be true. In either case in the final period when there is no use for further information, the consumer will purchase either one good or the other.

The shape of the EU_1 function is stated as propositions 4.2 and 4.3.

Proposition 4.2: The expected utility function EU_1 is convex in x_1 at low values of x_1, and concave at high values.

Proposition 4.3: If $\mu_o < p$ then

(i) EU_1 is everywhere a decreasing function of x_1.

or (ii) EU_1 has a maximiser x_1^*, in the interval $(0, \frac{M_1}{p})$

or (iii) EU_1 is decreasing in x_1 at $x_1 = 0$, but increasing at $x_1 = \frac{M_1}{p}$

if $\mu_o \geq p$ then

(iv) EU_1 is an increasing function of x_1.

These propositions can be illustrated diagrammatically:

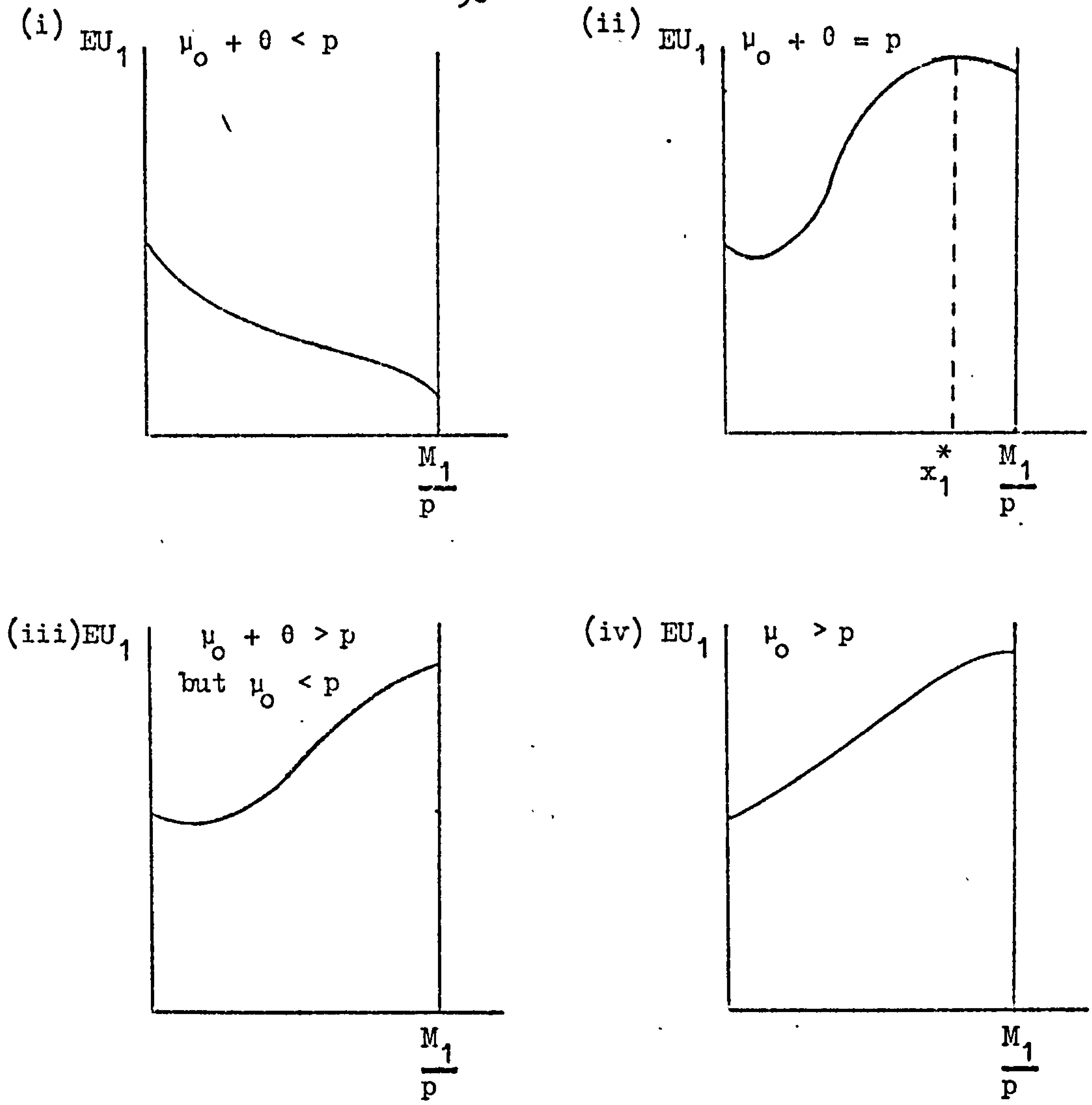


Diagram 4.2

The most interesting case is where the interior solution is defined. The optimal policy for the consumer is then to mix the consumption of the two goods in the first period. Whether the model yields an internal solution or not depends upon the values of the five parameters: p, M_1, M_2, μ_o and ϕ .

In order to simulate possible combinations of these parameters, (4.24) can be arranged in a more ameniable form.

Note that

$$\int_{-\infty}^{m^*} f(m)\ dm = 1 - \int_{m^*}^{\infty} f(m)\ dm$$

Substituting into equation (4.24), and rearranging

$$EV_2 = M_2 + \frac{M_2}{p} \cdot (\mu_o - p) \int_{m^*}^{\infty} f(m|0,1)dm + \frac{M_2}{p}\left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \frac{1}{(x_1+\phi)^{\frac{1}{2}}} \int_{m^*}^{\infty} mf(m|.)dm$$

But $\int_{m^*}^{\infty} mf(m|0,1)dm = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(m^*)^2}{2}}$

Substituting into equation (4.21), and assuming the rate of discount is zero:

$$EU_1 = M_1 + M_2 + (\mu_o - p)x_1 + \frac{M_2}{p}(\mu_o - p) \int_{m^*}^{\infty} f(m)dm + \frac{M_2}{p}\left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \frac{1}{(x_1+\phi)^{\frac{1}{2}}} \frac{e^{-\frac{(m^*)^2}{2}}}{\sqrt{2\pi}} \tag{4.28}$$

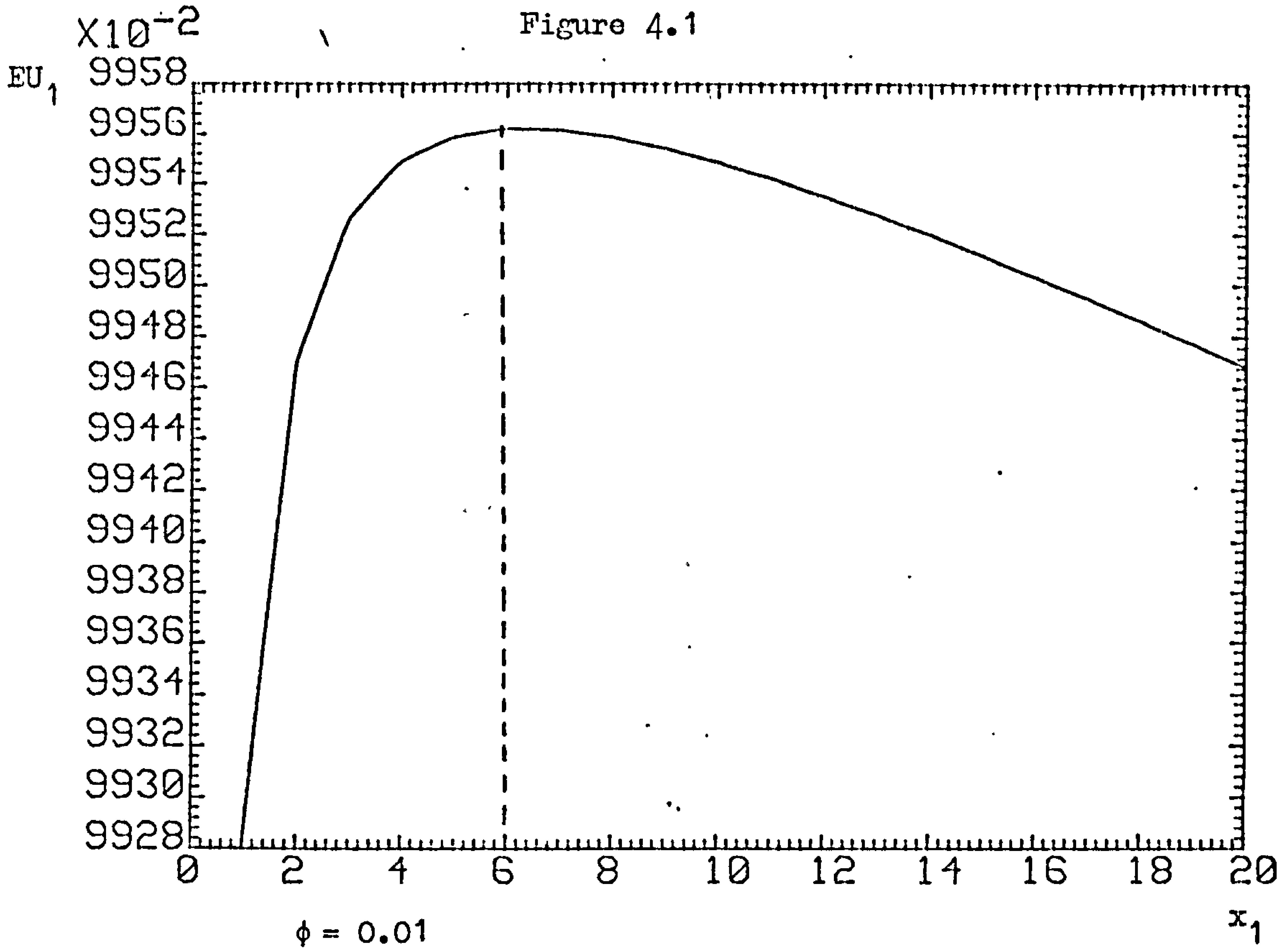
This function can be evaluated for specific values of the parameters.

5. Comparative Statics

Proposition 4.4: A consumer who maximises equation (4.21) will have a higher expected utility the more risky is the initial prior distribution of the unknown parameter. Risky being defined as a high initial variance, which is the inverse of the degree of precision.

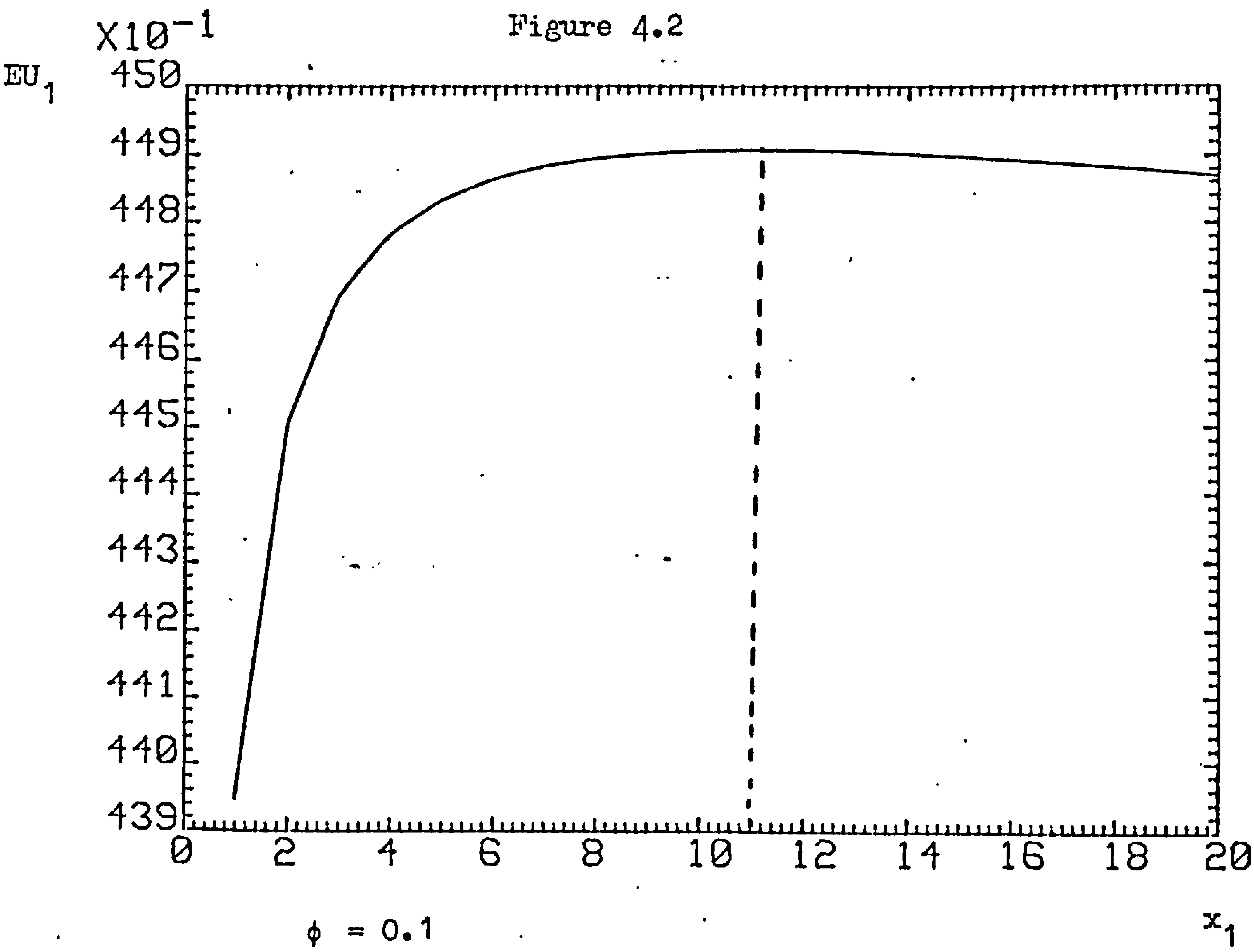
That is, increased variance increases expected utility. The reason for this result is that the more uncertain consumers are about their environment, the greater the expected utility of information. (See Rothschild, 1974a, p.691). The effect of a reduction in ϕ, on the diagrams above, will be to pivot the EU_1 function upwards around the point where $x_1 = 0$; since at $x_1 = 0$, ϕ does not enter the expected utility function.

Figure 4.1



$\phi = 0.01$

Figure 4.2



$\phi = 0.1$

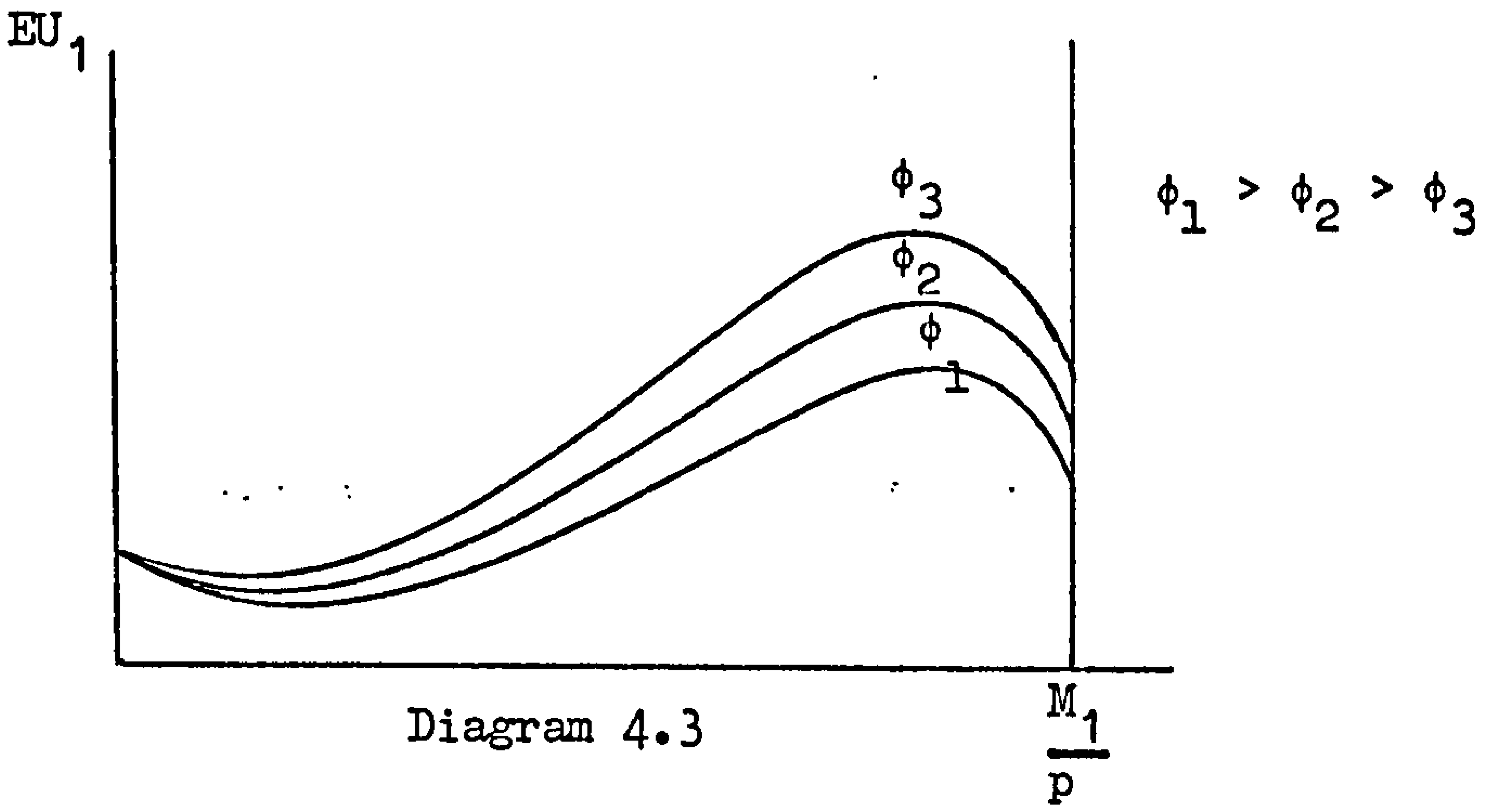


Diagram 4.3

Diagrame 4.3 and Table 1 illustrate proposition 4.4. As ϕ falls, expected utility increases.

Table 1

	p	μ_0	M_1	M_2	ϕ	x_1^*	EU_1 at $x_1 = 1$
(i)	0.5	0.49	10	10	1.0	19	27.48723
(ii)	0.5	0.49	10	10	0.1	11	43.94726
(iii)	0.5	0.49	10	10	0.01	6	99.28255
(iv)	0.5	0.49	10	10	0.001	4	272.0773
(v)	0.5	0.49	10	10	0.00001	0	2545.01
(vi)	1.2	1.19	10	10	0.01	4	53.02855

Figures 1 and 2, plot the values of expected utility as the purchase of the new good increases for the parameter values in (ii) and (iii).

The computer program used to generate these results is given in the Appendix to Chapter 8.

The reason that a decrease in ϕ increases expected utility, is due to the form of the expected utility function in (4.20). A decrease in ϕ increases the variance. Following a mean preserving spread, illustrated in diagram 4.4, the density function with the higher variance has an increased probability of higher values of ε_i. This is not correspondingly compensated by an increased probability of lower values, since the existence of the cut-off point $\bar{\varepsilon}^*$ guarantees a minimum level of utility M. Thus expected utility increases

Text cut off in original

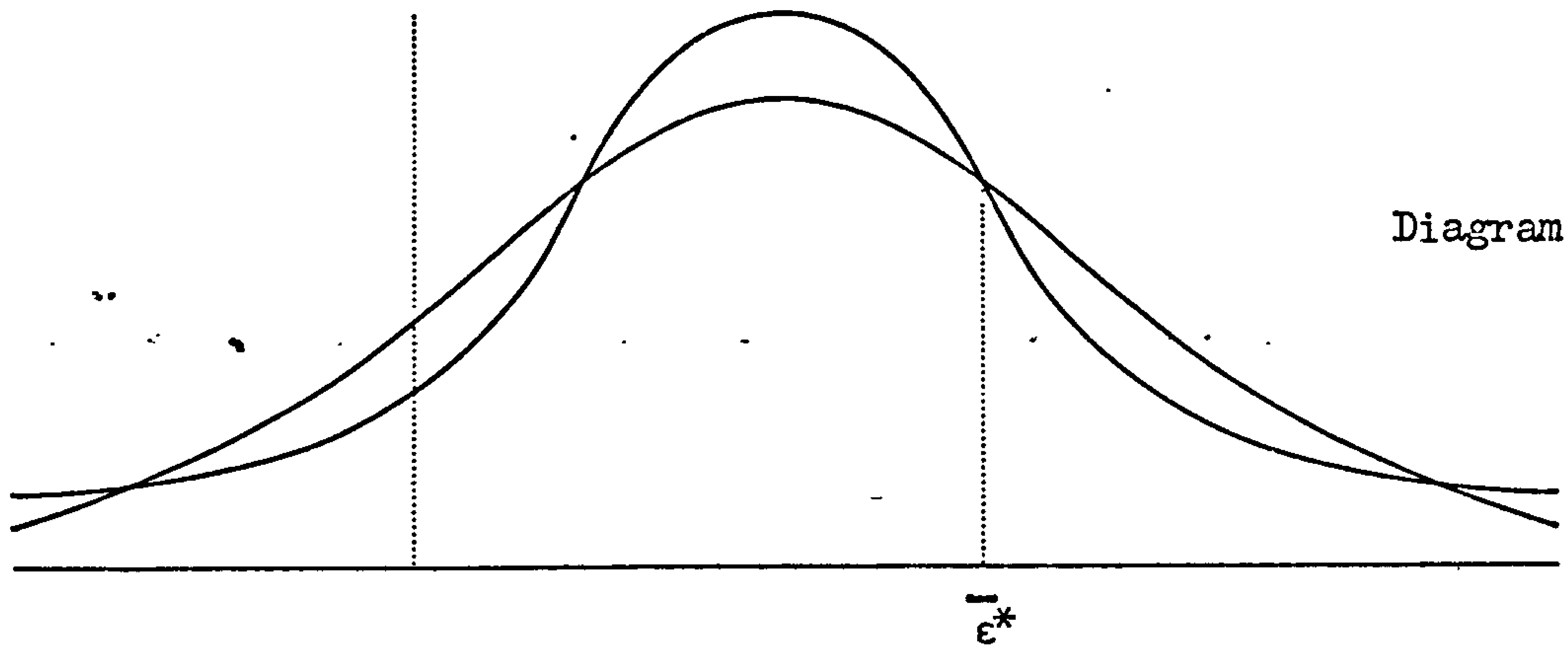


Diagram 4.4

Of interest is the situation where $\mu_o < p$ and for a sufficiently high value of ϕ, we would observe diagram 4.2i. Then as ϕ decreased, the curve would move upwards into a shape similar to diagram 4.2ii and finally to diagram 4.2iii. When the peak of the EU_1 function is equal to the EU_1 function at $x_1 = 0$, then the consumer is indifferent between $x_1 = 0$ and $x_1 = x_1^*$

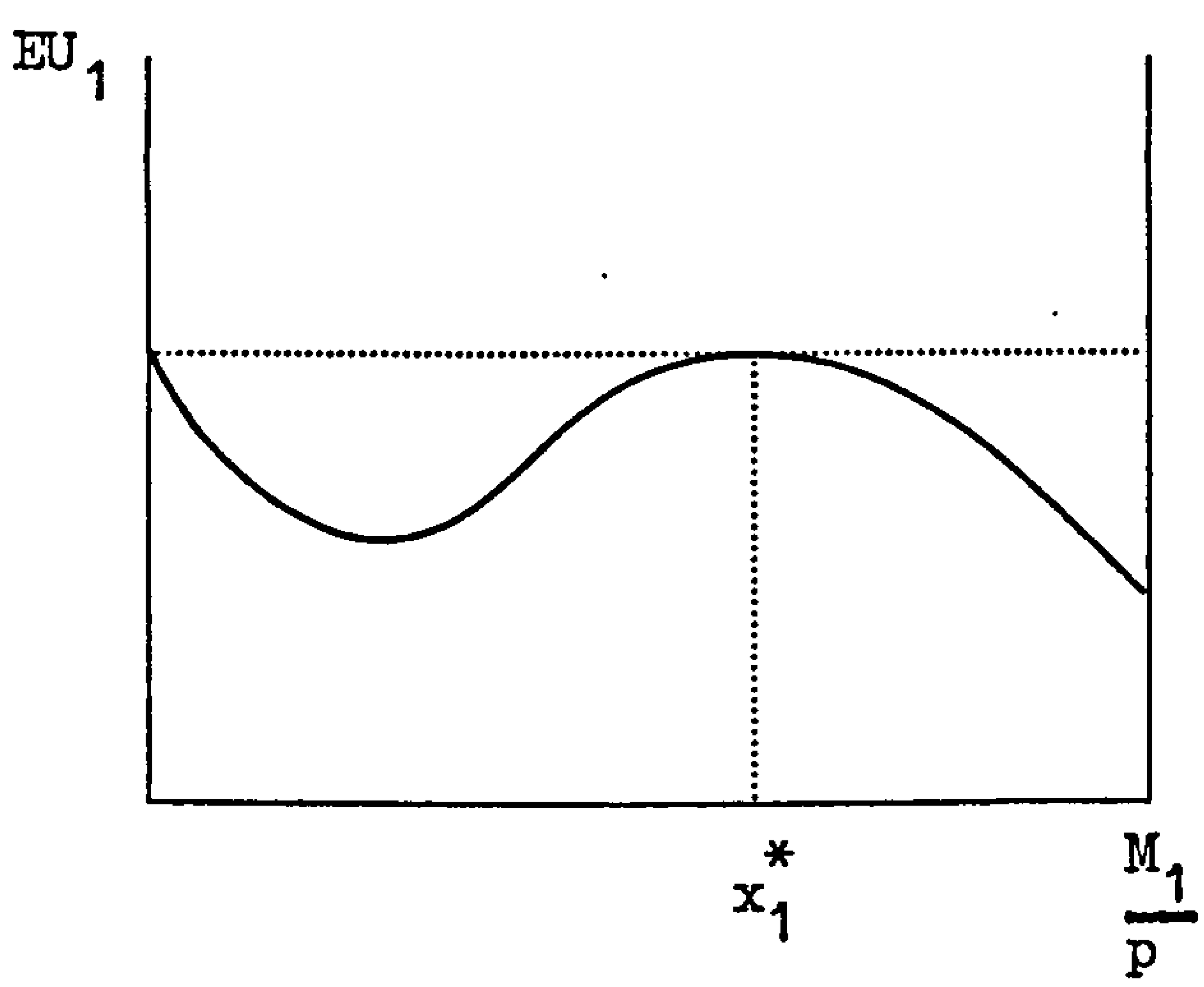


Diagram 4.5

Diagram 4.5 implies that for a small change in ϕ the optimal policy for the consumer is suddenly to switch from purchasing none of the new good to x_1^*. The consumer will not gradually increase its purchases of the new good as the degree of precision falls, but will move straight away to the optimal sample size: if it is optimal for the consumer to sample, it is optimal to sample optimally!

It is worthwhile contrasting the difference between an increase in x_1 and ϕ on the objective function, EU_1. EV_2 is increasing in x_1, but decreasing in ϕ, even though an increase in either parameter reduces the variance of the random variable in the second period. Proposition 4.4 shows that the higher is the initial variance the higher is expected utility, since the greater is $\frac{1}{\phi}$ the greater the returns from information. x_1 is the information component, not only does an increase in x_1 reduce the variance of the random variable in the next period, but puts a greater weight on the consumer's experience relative to its initial beliefs, as seen in equation (4.15).

Looking at Table 1 it can be seen that although an increase in the initial subjective variance increases expected utility, the increase in the variance results in a smaller optimal value of x_1^*.

Proposition 4.5: If $x_1^2 > (p - \mu_o)^2 \quad \phi[x_1^2 + 3\phi x_1 + 2\phi^2] + 2\phi x_1$ then x_1^* is monotonically increasing in ϕ.

This is illustrated in Fig.3, where $\mu_o = 0.49$, $p = 0.5$, $M = 80$ and ϕ is reduced from an initial value of 0.1 to 0.095. Consequently the optimal value of x_1 falls from 32 to 31 units.

Now consider the effect of a change in the mean of the unknown variable on the demand for the new good.

Proposition 4.6: x_1^* is monotonically non-decreasing in μ_o.

Figure 4.3

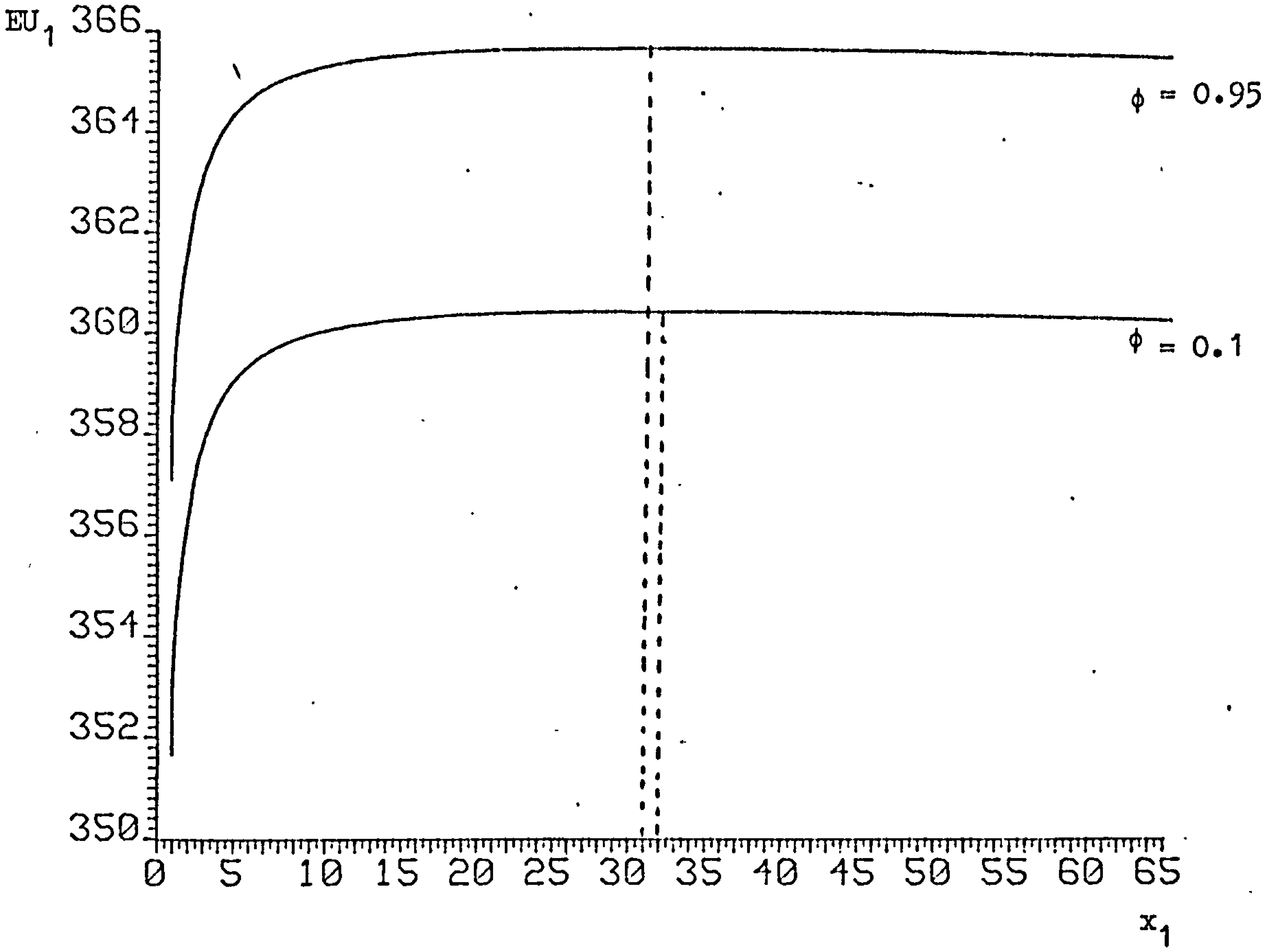
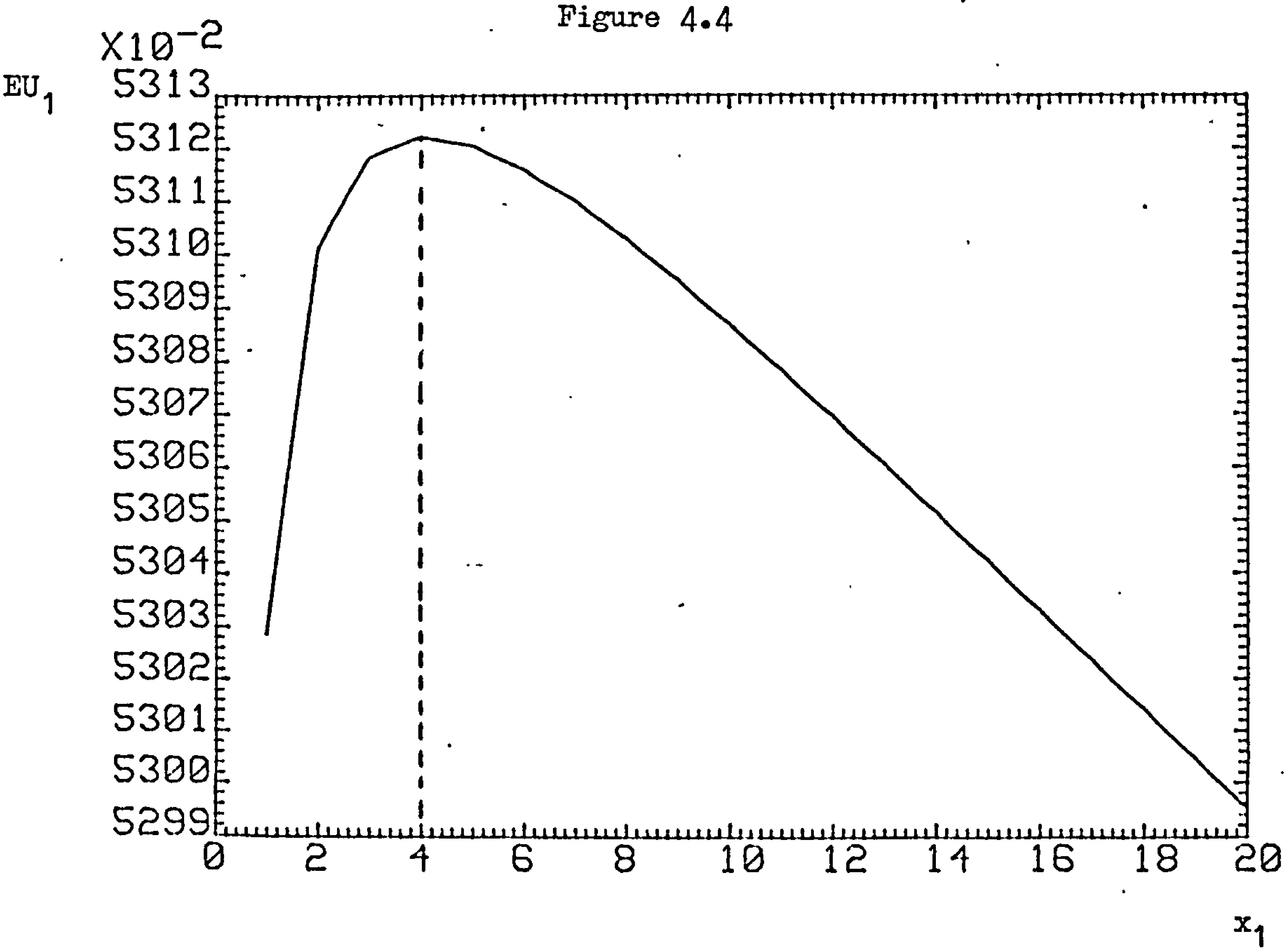


Figure 4.4



If the expected quality of the new good is higher, the consumer will demand more of x_1 both because it is a better buy and because he requires more information about it.

Proposition 4.7: x_1^* is monotonically non-decreasing in M_2

The richer is the consumer in the second period the more information he will demand now. This proposition implies that "information" is a normal good, in that an increase in income results in an increase in the demand for information.

Proposition 4.8: x_1^* is monotonically non-increasing in p .

The consumer purchases less of the new good if its price rises, because the good itself is more expensive and consequently the price of information has risen resulting in less demand.

Figure 4 plots case (vi) in Table 1. Both p and μ_o are increased with $p - \mu_o$ remaining the same. The result is a fall in expected utility and a fall in the optimal value of x_1^* .

From Table 1 we are able to make some generalisations about the restrictions on the parameter values necessary to generate an internal solution.

(i) p must be greater than μ_o , but $p - \mu_o$ must be small. If $\mu_o > p$ then we have the case of diagram 4.2iv and EU_1 is everywhere increasing in x_1 . Similarly if p is much larger than μ_o we have the case of diagram 4.2i . The absolute values of p and μ_o are not important, but their relative values are.

(ii) The value of ϕ should be less than unity, for the values of the other parameters in Table 1. If ϕ is too large, the expected utility function has the shape in 4.2i. If ϕ is too small, then since we are only dealing with integer values, the optimal value of x_1 is zero.

6. Flexible Budget Constraint

In this section we introduce a new budget constraint which is specified as

$$y_2 + px_2 = (1+i)\,[\, M - (y_1 + px_1)\,] \qquad (4.29)$$

The consumer is given a sum of money M at the beginning of the first period and he is allowed to spread this income over the two periods, any money not spent in the first period is invested at the market rate of interest which on account of perfect capital markets is assumed equal to the discount rate. The consumer must have spent all of his income by the end of the second period.

The consumer faces a new objective function

$$\max_{x_1, y_1} \quad EU_1' = y_1 + \mu_o x_1 + \frac{1}{1+i}\, EV_2' \qquad (4.30)$$

$$\text{where } EV_2' = (1+i)[M-(y_1+px_1)].\int_{\bar{\varepsilon}^*}^{\infty} \mu_{\bar{\varepsilon}}\, f(\bar{\varepsilon})d\bar{\varepsilon} + (1+i)\; M-(y_1+px_1)\int_{-\infty}^{\bar{\varepsilon}^*} f(\bar{\varepsilon})d\bar{\varepsilon} \quad (4.31)$$

Transforming the distribution of $\bar{\varepsilon}$ into a standard normal distribution 4.31 becomes

$$EV_2' = (1+i)[M-(y_1+ px_1)]\; \{\frac{1}{p}\int_{m^*}^{\infty} [\mu_o + (\frac{x_1}{\phi})^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}}]\; f(m)dm + \int_{-\infty}^{m^*} f(m)dm\} \quad (4.32)$$

where as previously m* is defined by equation (4.23).

The consumer now has two decision variables in the first period: x_1 and y_1. He does not have to spend all of his budget by the end of the first period; in fact the less he spends in the first period the more he will have left to spend when he is more certain of the distribution of product quality. It can be shown, rather surprisingly, that in this case the consumer will not purchase any of the old good in the first period.

Proposition 4.9: If the expected utility function is given by (4.30) then $y_1^* = 0$.

This proposition does not mean that the consumer will never purchase the old good, only that he will not purchase it in the first period: it will not be optimal to mix the two goods, which can be contrasted with proposition 4.3ii. On the other hand, proposition 4.9 does not mean that the consumer will always sample, it may also be the case that $x_1 = 0$, in which case in the second period the consumer will spend his entire budget on the old good, which is essentially a one period solution, i.e. $y_2 = M$. If the consumer decides to purchase the old good, then he will choose that good for ever, since by purchasing the old good he can gain no new information about the quality of the new good which may persuade him to switch to the new good.[6]

In a sense, time has become the consumer's decision variable; since he is living in a world of perfect markets, the consumer is able to experiment with the new good for as long as is optimal and then make a decision between the new and old goods.

It can also be shown that it will never be optimal for the consumer to spend his entire budget sampling.

Proposition 4.10: If the consumer's objective function is given by (4.30) then $x_1^* < \frac{M}{p}$

This proposition states that it will never be optimal for the consumer to spend his entire budget sampling, since he will have no money left to benefit from the information he has gained.

The shape of the EU_1' function is drawn in diagrams 4.6 and 4.7 .

Diagram 4.6

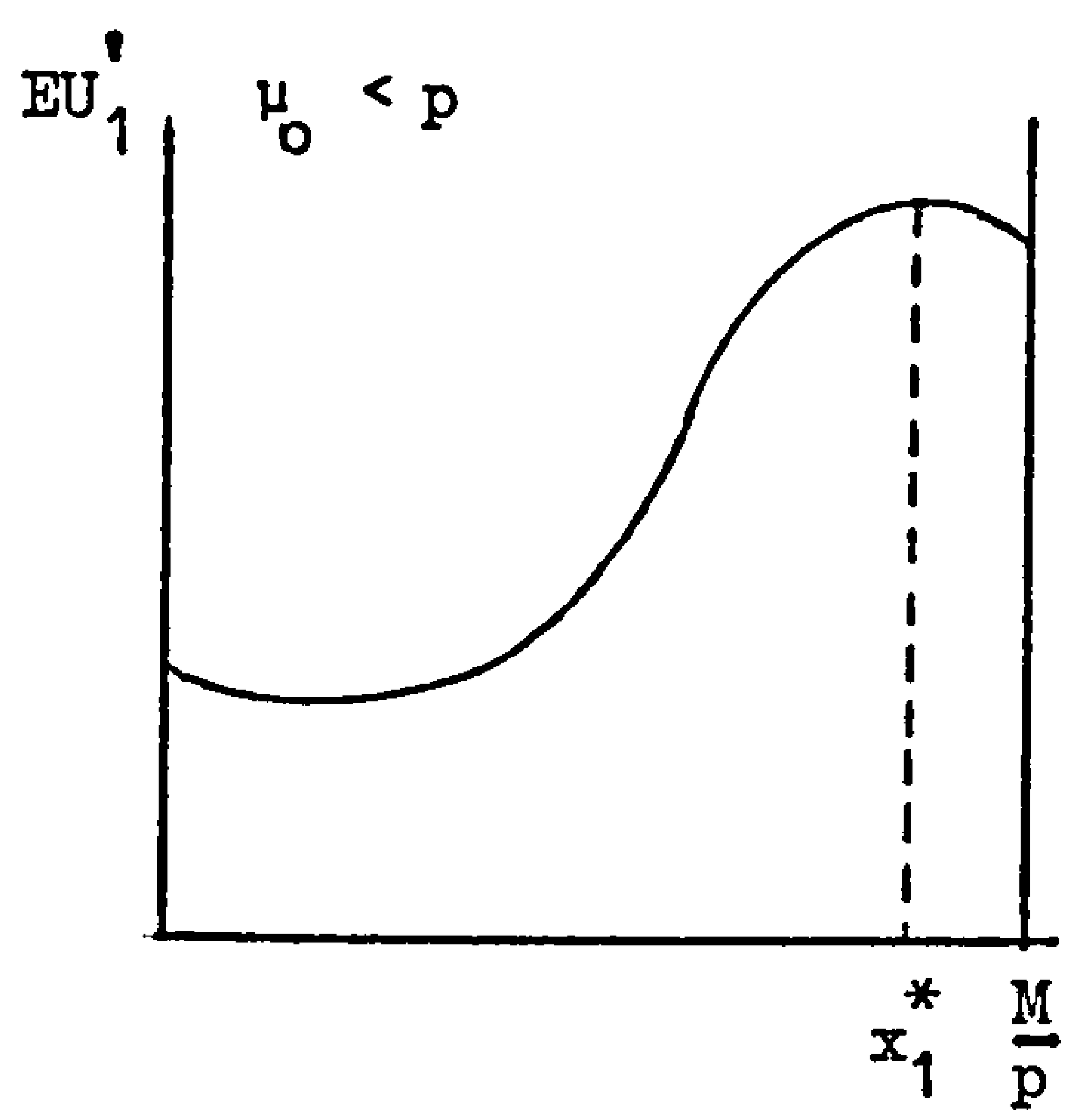
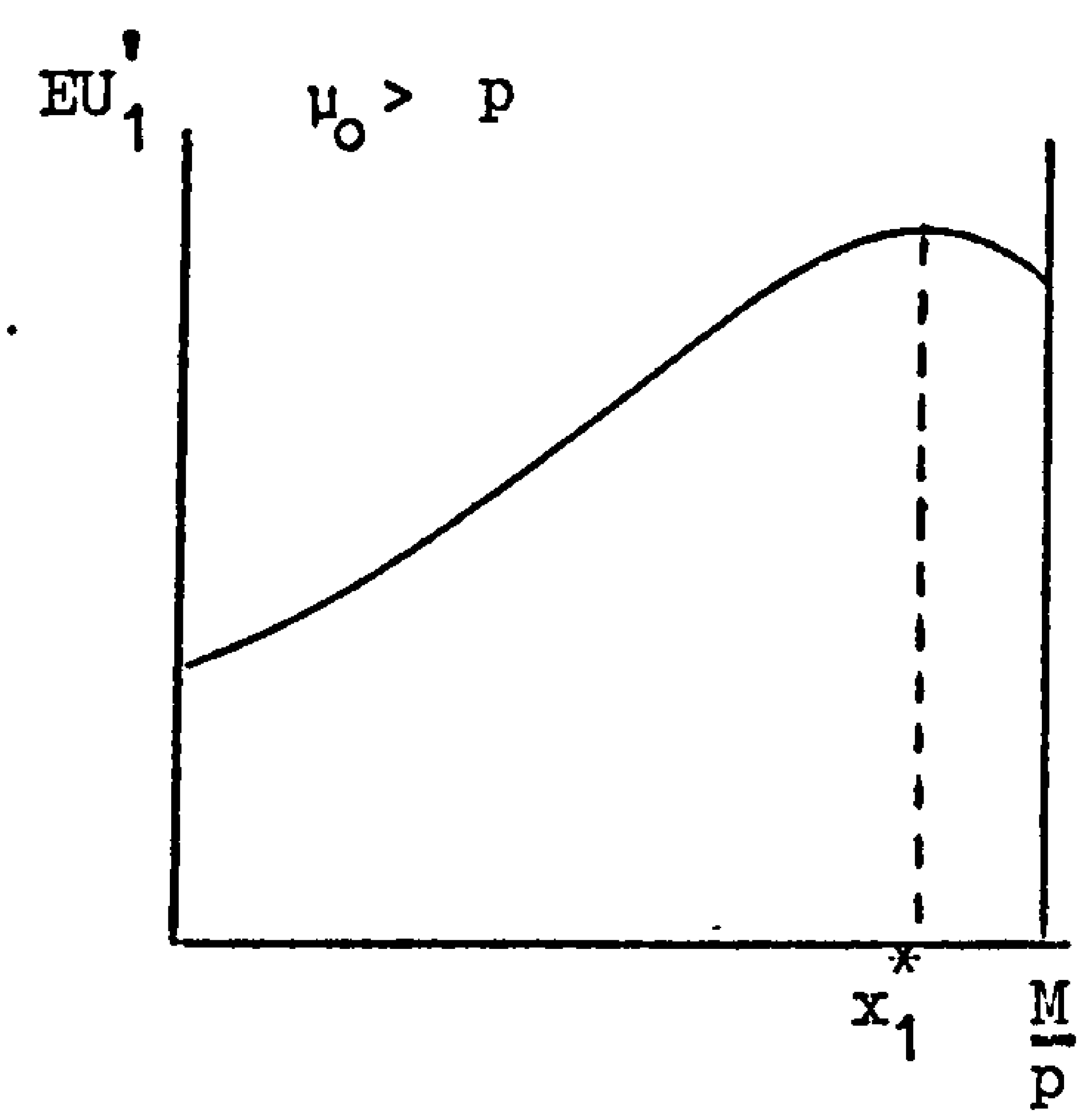


Diagram 4.7



In diagram 4.7 when $\mu_o > p$, it will always be optimal for the consumer to sample. In diagram 4.6 when $\mu_o < p$, $x_1^* > 0$ depends upon the relative values of the parameters.

Equations (4.30) and (4.32) can be rearranged into a form which can be calculated on a computer in the same way as equation (4.28) was derived.

In this case we obtain

$$EU_1' = M + (\mu_o - p)x_1 + \frac{(M-px_1)}{p}\left\{(\mu_o - p)\int_{m^*}^{\infty} f(m)dm + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{1}{(x_1+\phi)^{\frac{1}{2}}} \cdot \frac{e^{-(\frac{m^*}{2})}}{\sqrt{2\pi}}\right\} \quad (4.33)$$

Tables 2, 3 and 4, and figures 5 and 6, show the results of evaluating EU_1' and the optimal value of x_1, for various values of the parameters. As can be seen from Tables 2 and 3, in line with proposition 4.4, a reduction in the degree of precision increases the level of expected utility; also proposition 4.5 carries over to the flexible budget case: a reduction in ϕ reduces the optimal value of x_1. Figure 5 illustrates Table 3.

By comparing Table 2 with Table 3, we can see that proposition 4.7 holds: an increase in income increases the value of x_1^*.

Table 4 and Figure 6, illustrate the case of $\mu_o > p$, which with the previous non-flexible budget constraint resulted in a corner solution with the consumer purchasing all of the new good. In this case, as proposition 4.10 stated, the consumer will never spend all of his income sampling.

Table 4 shows that although the consumer will purchase more of the new good as μ_o increases, it is still true that $x_1^* < \frac{M}{p}$.

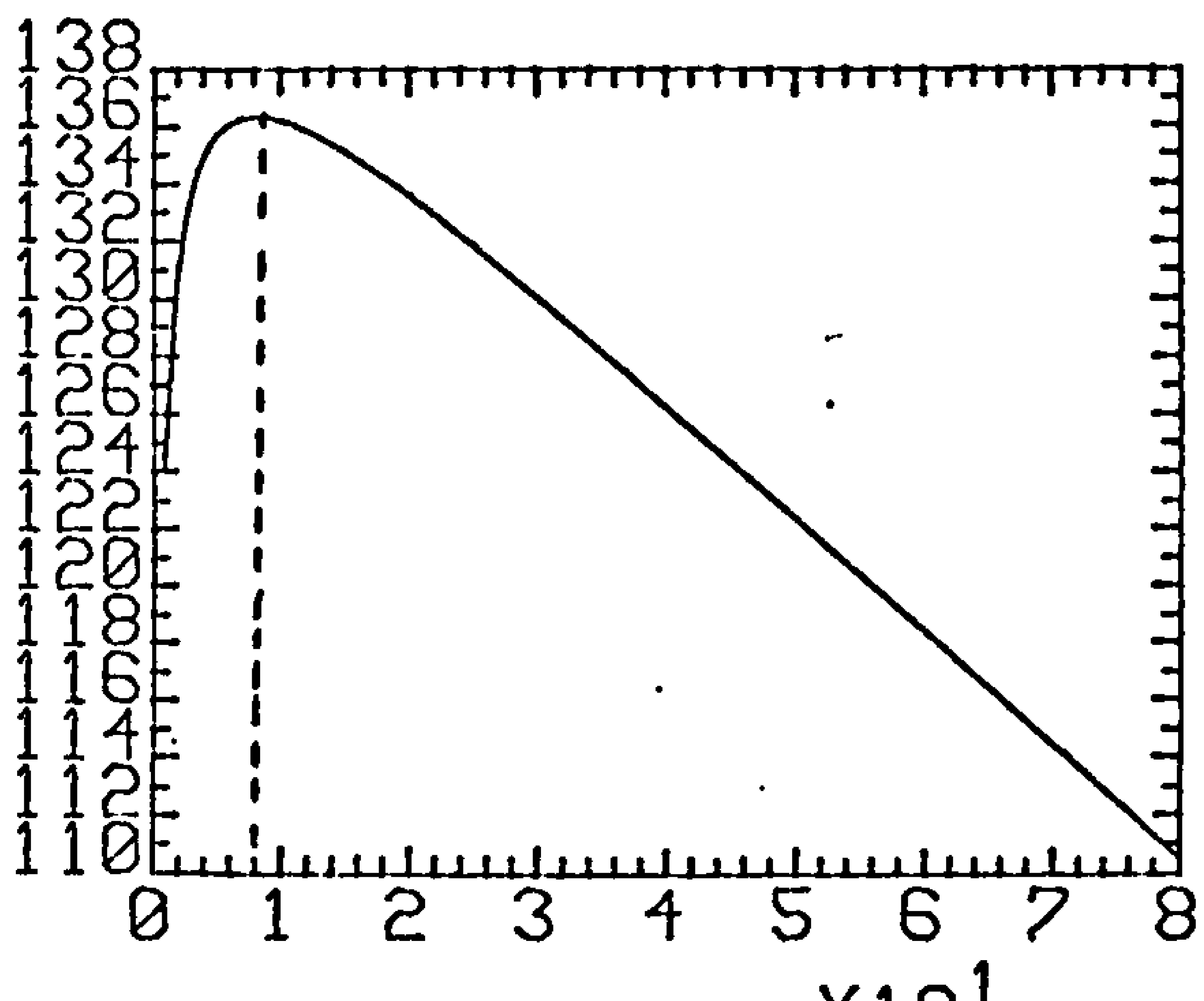
<u>Table 2</u>

	μ_o	p	M	ϕ	x_1^*	EU_1' at $x_1 = 1$
(i)	0.49	0.5	20	10.0	8	21.2866
(ii)	0.49	0.5	20	2.0	5	26.1487
(iii)	0.49	0.5	20	1.0	4	32.1688
(iv)	0.49	0.5	20	0.5	3	37.7613

<u>Table 3</u>

(v)	0.49	0.5	80	50.0	33	80.6079
(vi)	0.49	0.5	80	25.0	28	81.7634
(vii)	0.49	0.5	80	5.0	16	90.7933
(viii)	0.49	0.5	80	1.0	8	124.0526

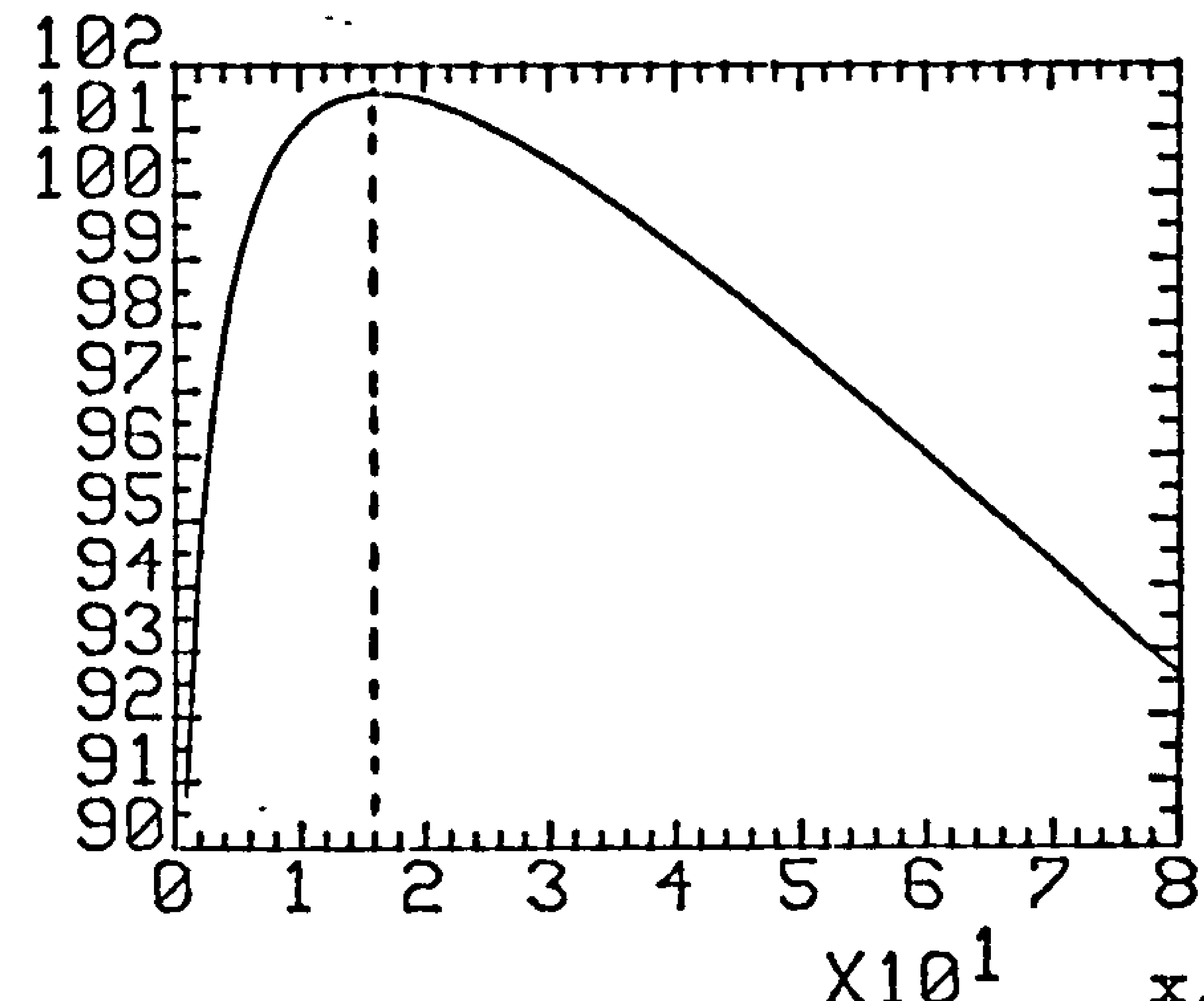
Figure 4.5(viii)



$\phi = 1.0$

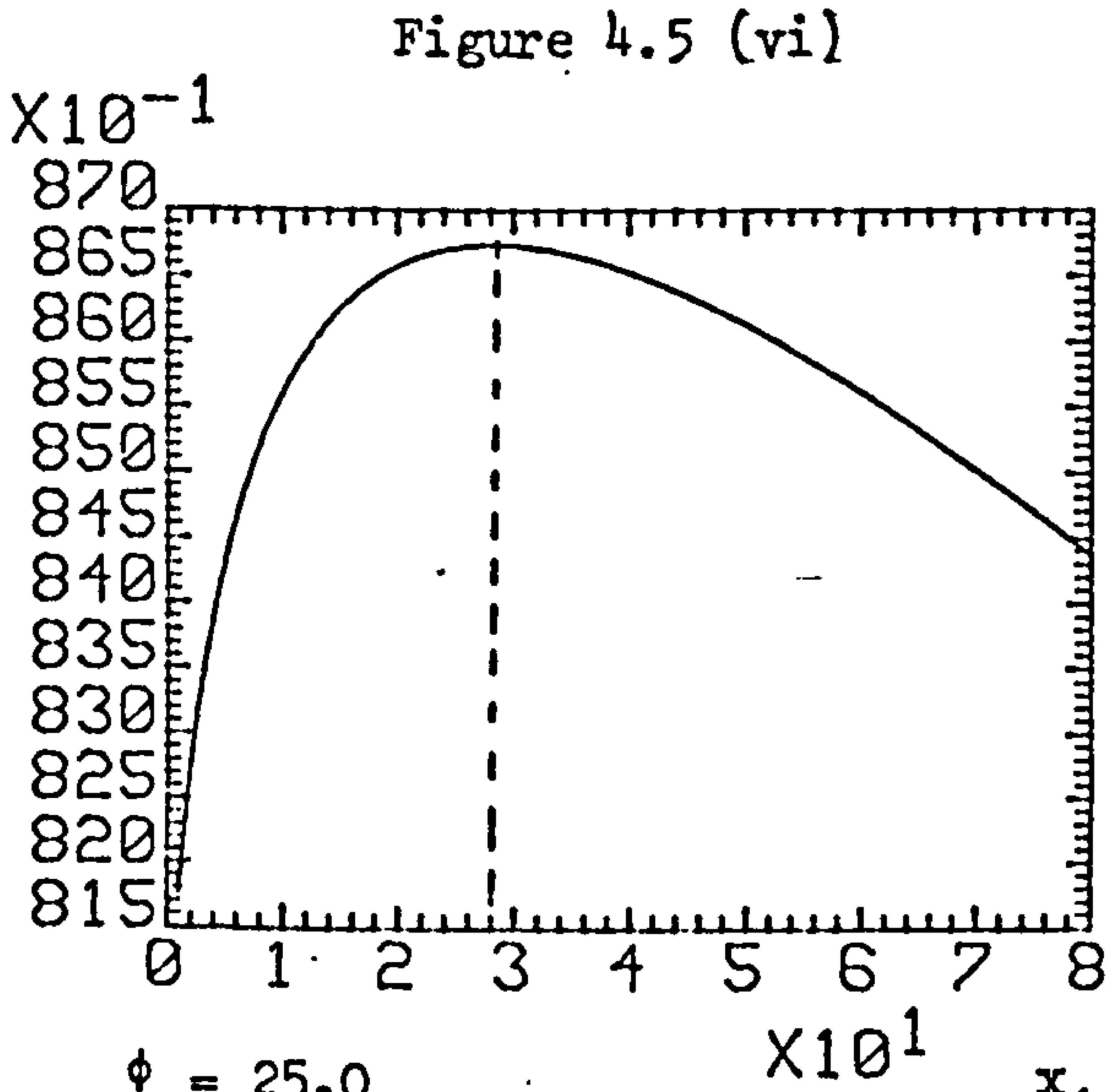
EU$_1$

Figure 4.5(vii)



$\phi = 5.0$

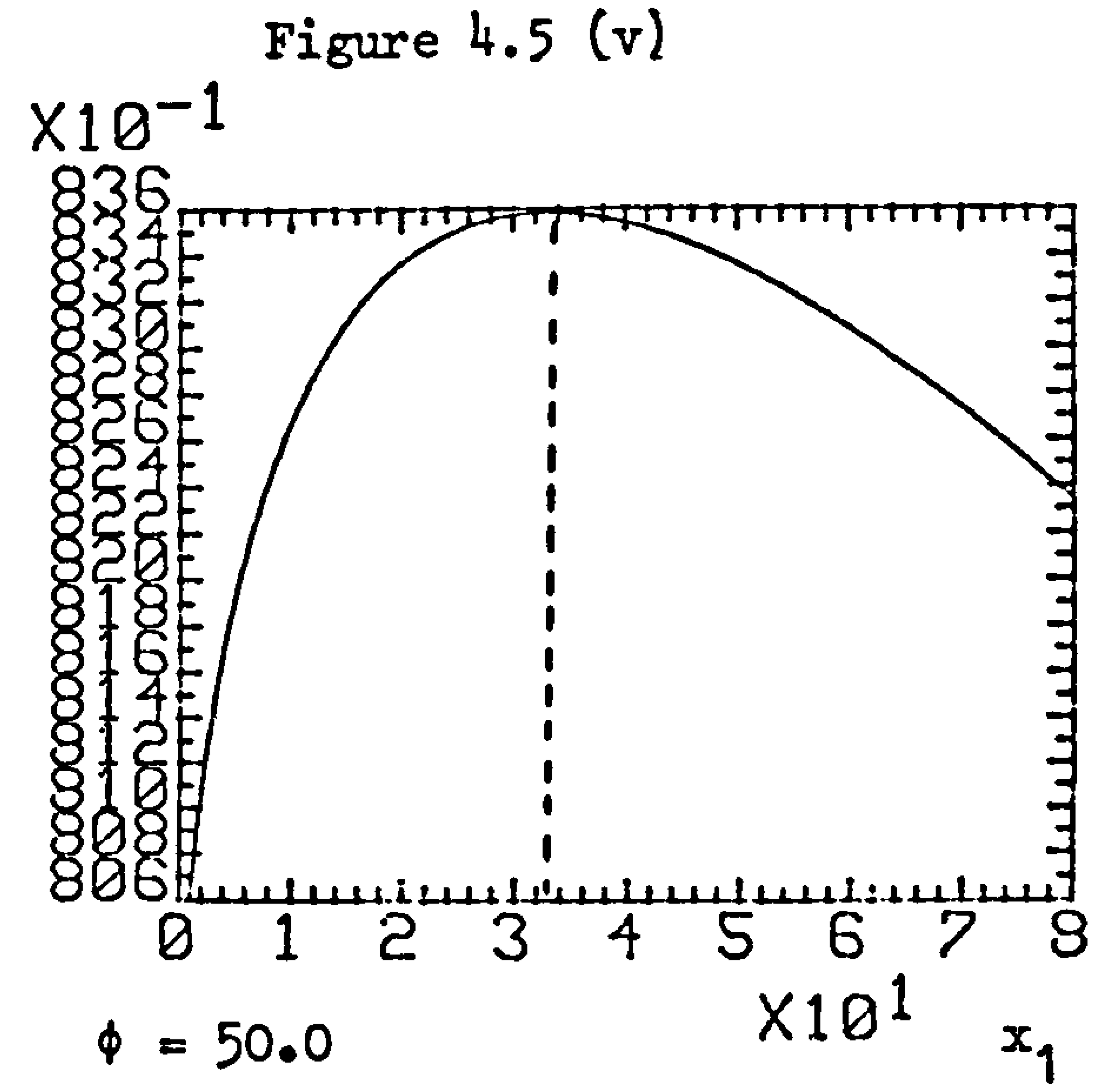
Figure 4.5 (vi)



$\phi = 25.0$

EU$_1$

Figure 4.5 (v)



$\phi = 50.0$

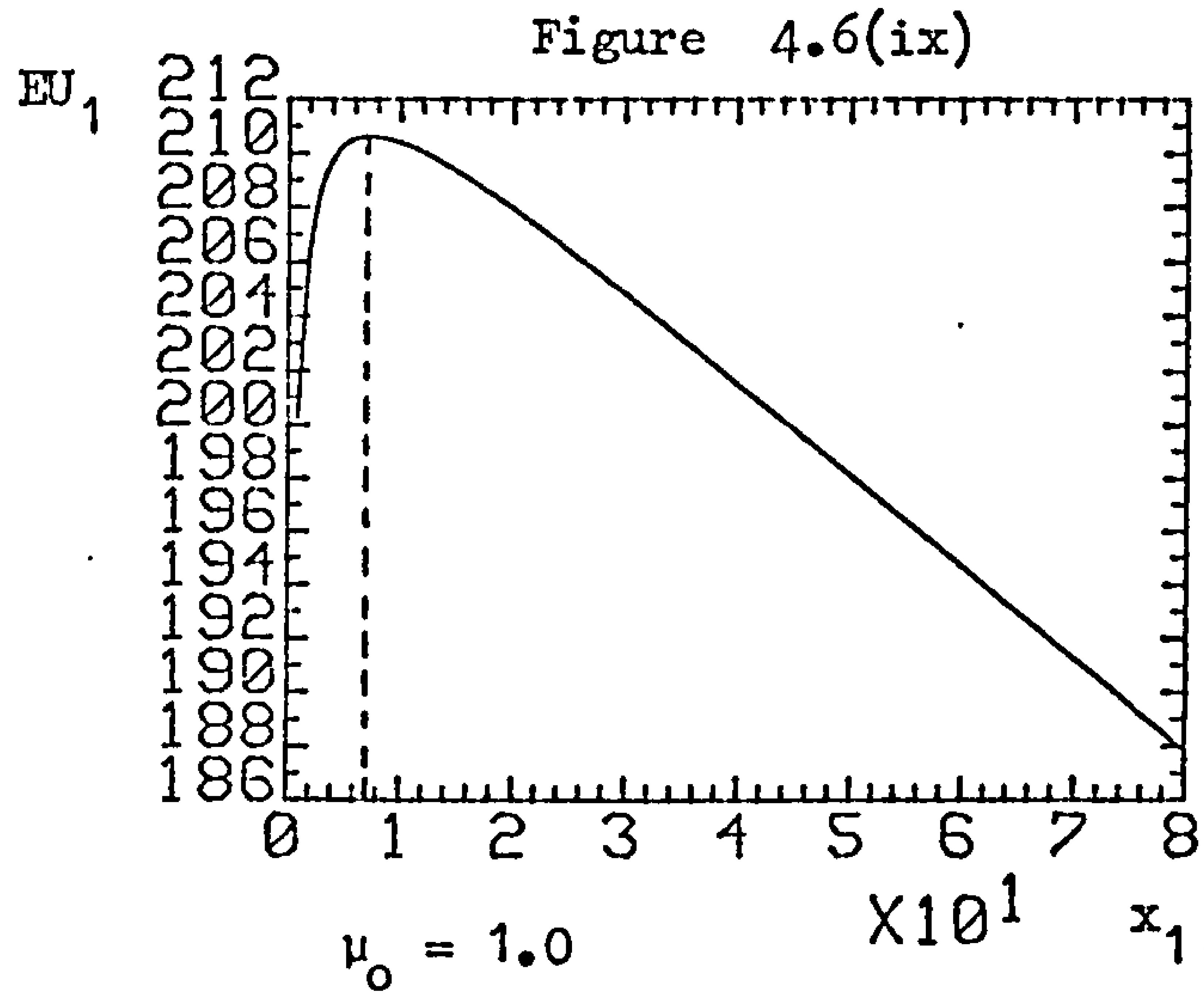


Figure 4.6(ix)

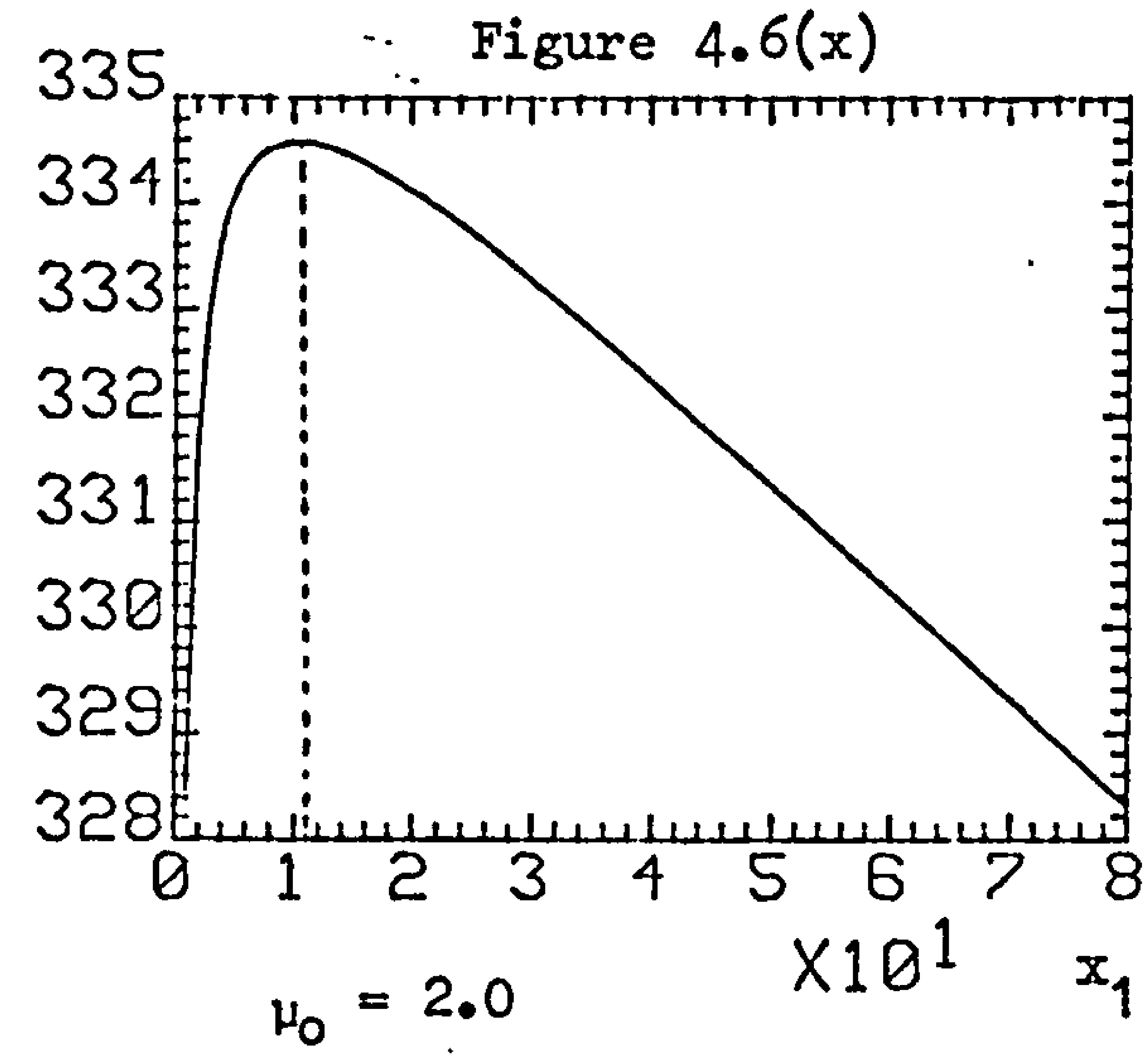


Figure 4.6(x)

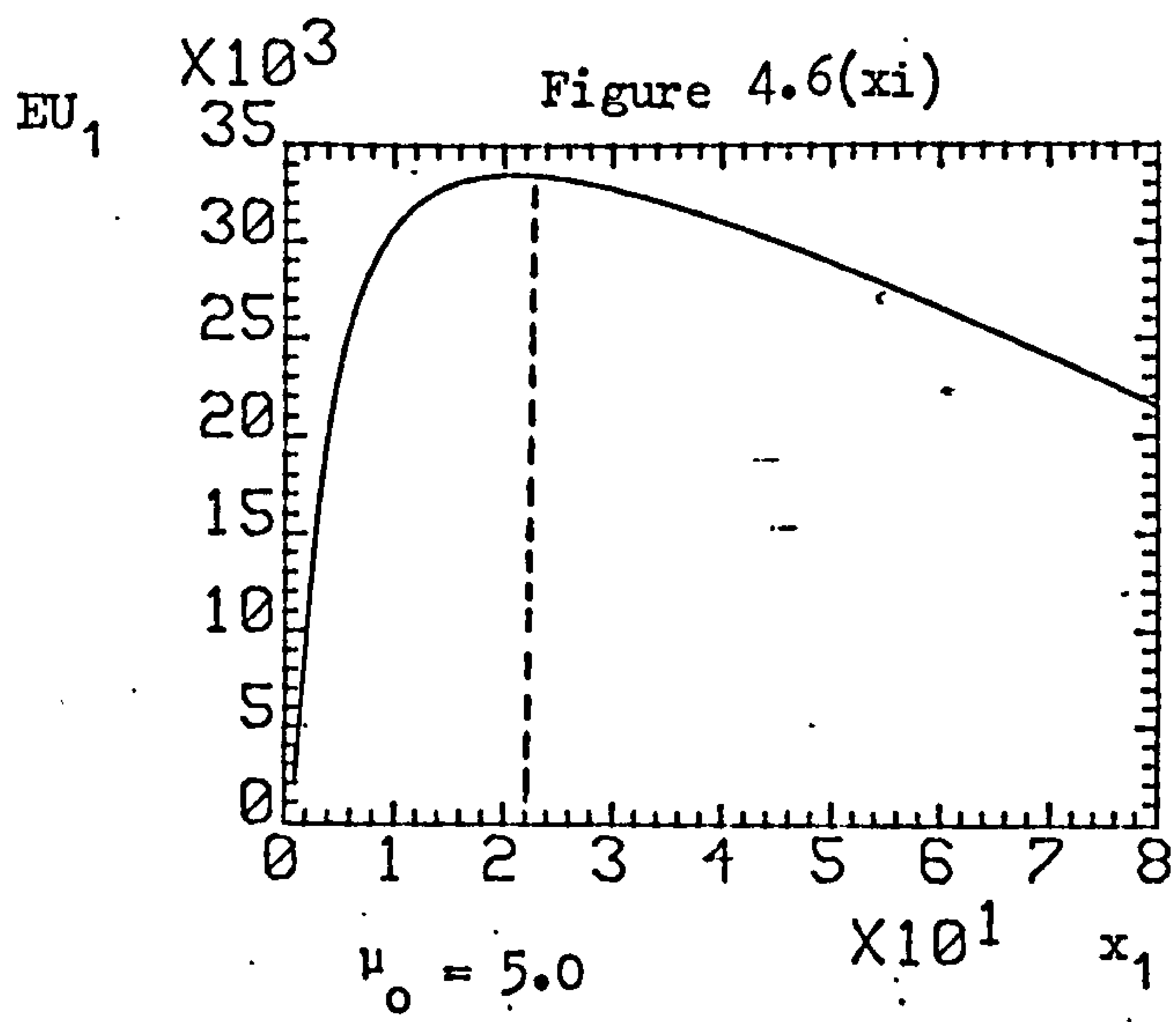


Figure 4.6(xi)

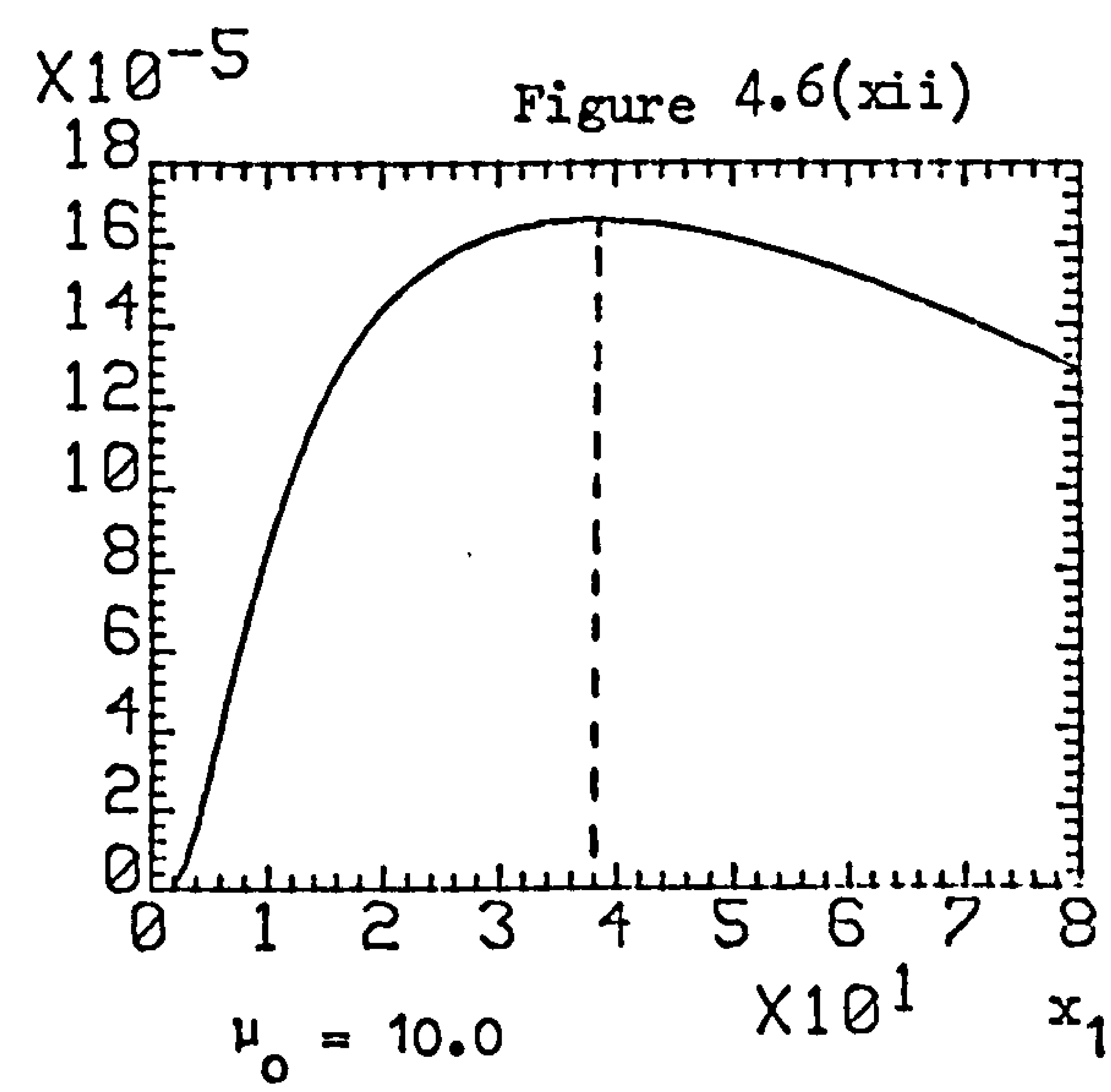


Figure 4.6(xii)

Table 4

	μ_o	p	M	ϕ	x_1^*	EU_1' at $x_1 = 1$
(ix)	1.0	0.5	80	0.5	7	200.2562
(x)	2.0	0.5	80	0.5	11	328.3757
(xi)	5.0	0.5	80	0.5	21	800.0021
(xii)	10.0	0.5	80	0.5	38	1600.0000

When the difference between quality and price is very large, such as in line (xii), the expected value of information is low, and in fact the maximum value function for a single period utility function : $\mu_o \cdot \frac{M}{p} = 1600.0$. Thus the consumer is virtually indifferent as to the quantity of x_1^* consumed. But the tiny expected value of information means that an internal solution is defined.

We shall now state and prove a theorem that uses the flexible budget constraint to contradict Theorem 4.1.

Theorem 4.2: If x_1^o is the value of x_1 which maximises the expected value of future utility in the non-adaptive case, and x_1^* is the optimal value in adaptive case; then if the rate of time preference is greater than the market rate of interest by a small amount τ, where

$$0 < \tau < \frac{1+i}{\mu_o} \int_{-\infty}^{m^*} \left[\mu_o - p + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \frac{m}{(\phi + x_1)^{\frac{1}{2}}} \right] f(m)dm$$

and if $\mu_o > p(1 + \frac{\tau}{1+i})$

then $x_1^o > x_1^*$

Proof: We suppose that in the non-adaptive case the consumer uses his prior density function in each period. Firstly consider the situation where the rate of time preference is the same as the market rate of interest, and suppose $\mu_o > p$. Then a consumer with a non-adaptive distribution in the final period will maximise

$$EU_2^o = \max_{x_2} \{ (1+i)[M - (y_1 + px_1) - px_2 + \mu_o x_2] \}$$

F.O.C.

$$\frac{dEU_2^o}{dx_2} = -p + \mu_o > 0$$

By assumption $p < \mu_o$, so $x_2 = \dfrac{(1+i)[M-(y_1 + px_1)]}{p}$

$$\text{So} \quad EV_2^o = \mu_o \frac{[M - (y_1 + px_1)] (1+i)}{p} \tag{4.34}$$

Then we may write consumer's objective function as

$$\max_{y_1, x_1} \quad EU_1^o = y_1 + \mu_o x_1 + \frac{1}{1+i} \; EV_2^o \tag{4.35}$$

$$\text{s.t.} \quad M \geq y_1 + px_1$$

F.O.C.'s yield:

$$\frac{dEU_1^o}{dy_1} = 1 - \frac{\mu_o}{p} < 0 \rightarrow y_1 = 0 \tag{4.36}$$

$$\frac{dEU_1^o}{dx_1} = \mu_o - \mu_o = 0 \tag{4.37}$$

Since the expected quality of the new good is greater than its relative price, the consumer will only purchase the new good. But equation (4.37) shows that how the consumer divides his purchases of the new good over the two periods is indeterminate.

The optimal solution can be made determinate by specifying a difference between the consumer's rate of time preference and the market rate of interest. Let the latter remain at i, but the former becomes $i + \tau$. This anomoly is taken to represent an anomoly in the capital market. Then (4.36) becomes:

$$\frac{dEU_1^o}{dy_1} = 1 - \frac{1+i}{1+i+\tau} \cdot \frac{\mu_o}{p} < 0$$

since by assumption $\mu_o > p\left(1 + \frac{\tau}{1+i}\right)$, so

the consumer still purchases none of the old good. Equation (4.37) becomes:

$$\frac{dEU_1^o}{dx_1} = \mu_o - \frac{1+i}{1+i+\tau} \cdot \mu_o > 0 ,$$

so the consumer spends his entire budget on the new good in the first period

$$x_1^o = \frac{M}{p} .$$

We know from proposition 4.10, that $x_1^* < \frac{M}{p}$, but is this conclusion altered by the introduction of a rate of time preference $(i+\tau)$? Making use of A12, we have

$$\left.\frac{dEU_1'}{dx_1}\right|_{x_1=\frac{M}{p}} = \mu_o\left[1 - \frac{1+i}{1+i+\tau}\right] + \frac{1+i}{1+i+\tau} \cdot \int_{-\infty}^{m^*}\left[\mu_o - p + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right] f(m)\,dm$$

$$= \mu_o \cdot \frac{\tau}{1+i+\tau} + \frac{1+i}{1+i+\tau} \cdot \int_{-\infty}^{m^*}\left[\mu_o - p + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right] f(m)\,dm$$

So $$\left.\frac{dEU_1'}{dx_1}\right|_{x_1=\frac{M}{p}} < 0 \quad \text{if} \quad \tau < \frac{1+i}{\mu_o} \int_{-\infty}^{m^*}\left[\mu_o - p + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right] f(m)\,dm$$

$$x_1^* < \frac{M}{p}$$

Q. E. D.

Provided $\grave{\tau}$ is small enough it will not affect the conclusions of proposition 4.10. The consumer's preferences dictate that with a non-adaptive distribution, he should spend his entire income on the new good in the first period. With the adaptive distribution, the fact that the consumer has income for use in the second period means that it will never be optimal to spend his entire income sampling.

Suppose that in order to encourage the consumer to purchase the new good, as a marketing device, the firm sets the price in the first period lower than the price in the second, and it is assumed that the consumer recognises this price differential exists.

The consumer's objective becomes:

$$\max EU_1'' = \mu_o x_1 + \frac{1}{1+i} EV_2'' \tag{4.38}$$

$$\text{s.t.} \quad M \geq p_1 x_1$$

where
$$EV_2'' = (1+i)\,[M-p_1x_1]\left\{\frac{1}{p_2}\int_{m^*}^{\infty}\left[\mu_o+\left(\frac{x_1}{\phi}\right)^{\frac{1}{2}}\frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right]f(m)\,dm+\int_{-\infty}^{m^*}f(m)\,dm\right\} \tag{4.39}$$

and
$$m^* = (p_2-\mu_o)\,(\phi+x_1)^{\frac{1}{2}}\left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \tag{4.40}$$

Then we can state the following propositions

Proposition 4.11: If the consumer's objective function is given by (4.38), where $p_1 < p_2$ and $p_1 < \mu_o$, then for sufficiently high values of m^*,

$$x_1^* = \frac{M}{p_1}$$

The prospect of an increase in price in the second period, coupled with an expectation that product quality is higher than the current price p_1, means that the consumer will not bother making use of information acquisition but will spend his entire budget on the new good in the first period. This proposition can be contrasted with Proposition 4.10.

Proposition 4.12: x_1^* is monotonically increasing or decreasing in p_2 as

$$\mu_o \gtrless \frac{p_1 F(m^*) + p_2 \dfrac{(M-p_1x_1)}{(\phi+x_1)^{\frac{1}{2}}} \cdot \left(\dfrac{\phi}{x_1}\right)^{3/2} \cdot f(m^*)}{1 + \dfrac{(M-p_1x_1)}{2} \cdot \left(\dfrac{\phi}{x_1}\right)^{3/2} \cdot \dfrac{f(m^*)}{(\phi+x_1)^{\frac{1}{2}}}}$$

Whether the effect is positive or negative depends upon whether the new good is expected to be of high or low quality.

For example if $\mu_o > 2p_2$ and $\mu_o > p_1F(m^*)$ then it can be seen from A14, that

$\frac{dx_1^*}{dp_2} > 0$. This is because if high quality is expected, then the consumer expects to purchase the new good after sampling, thus a rise in price in the second period causes the consumer to purchase more of the new good in the first period. The rise in p_2 induces inter-temporal substitution, with the consumer purchasing more of the new good now instead of purchasing it at a higher price next period.

If $\mu_o < p_1 F(m^*)$ and $\mu_o < 2p_2$, then it can be seen that $\frac{dx_1^*}{dp_2} < 0$. If the quality is expected to be low, then the consumer is unlikely to purchase the new good at any stage. A rise in price in either of the periods will consequently have a normal negative effect.

7. The demand for complementary goods with adjustment costs

So far, utility has been specified as a function of two goods x and y. But now suppose we introduce a third good s, price p^s, which is complementary to the first two goods and must be consumed in conjunction with goods x and y. Utility in any period t can be written

$$u_t = y_t + \tilde{\varepsilon}_t x_t$$

where
$$M_t \geq y_t + px_t + p^s(s_t^x + s_t^y) \tag{4.41}$$

and
$$\left.\begin{aligned} s_t^y &\geq ay_t \\ s_t^x &\geq bx_t \end{aligned}\right\} \tag{4.42}$$

s_t^y is the amount of good s which must be consumed in fixed proportion a, with good y. s_t^x is the amount of good s which must be consumed in fixed proportion b with good x, and $s_t = s_t^y + s_t^x$. We suppose that $a > b$, that is the new good needs less of good s to complement it than does the old good. This is to give some credence to the new good being a technical improvement.

The problem is complicated by the existence of adjustment costs, with respect to changes in the purchases of good s between periods. It is assumed that the household technology is constructed for a particular size of good s.[7] Any change in the quantity of s consumed means that the consumer must adjust the household technology. If the consumer purchases one extra unit of s, the adjustment cost is d, and if he purchases one less he incurs a cost of e.

An example of this type of technology would be the relationship between type of washing powder and size of washing machine. Suppose a new type of washing powder is developed, which needs a much smaller quantity of water, and hence smaller machine, to wash clothes in. If the consumer wishes to change the technology, he must buy a smaller washing machine.

If s_1 and s_2 are the quantities of good s purchased in periods one and two, then if $s_1 > s_2$, the adjustment cost is $e(s_1 - s_2)$; if $s_2 > s_1$, the total adjustment cost is $d(s_2 - s_1)$.

Furthermore we suppose that the household technology relationships in (4.42) are binding in the final period, since in the final period the consumer is in long run equilibrium, then substituting for y_2 in the budget constraint

$$s_2 = \frac{a}{1+ap^s} \left[M_2 - (p + bp^s)x_2 \right] + bx_2 \qquad (4.43)$$

So the adjustment costs become:

$$d \left\{ \frac{a}{1+ap^s} \left[M_2 - (p + bp^s)\, x_2 \right] + bx_2 - s_1 \right\} \quad \text{if} \quad s_2 > s_1 \qquad (4.44)$$

$$\text{and} \quad e \left\{ s_1 - \frac{a}{1+ap^s} \left[M_2 - (p + bp^s)\, x_2 \right] - bx_2 \right\} \quad \text{if} \quad s_1 > s_2 \qquad (4.45)$$

Then in the second period the consumer faces the following objective

$$\max_{x_2} EU_2 = \frac{M_2}{1+ap^s} - \frac{(p + bp^s)}{(1 + ap^s)}\, x_2 + \mu_{\varepsilon} x_2 - AC \qquad (4.46)$$

where AC are the adjustment costs specified in equations (4.44) and (4.45)

F.O.C.

$$\frac{dEU_2}{dx_2} = -\frac{(p + bp^s)}{(1 + ap^s)} + \underset{\varepsilon}{\mu} \begin{cases} -e\left[\frac{a}{1+ap^s}\cdot (p + bp^s) - b\right] & \text{if } s_1 > s_2 \\ +d\left[\frac{a}{1+ap^s}\cdot (p + bp^s) - b\right] & \text{if } s_2 > s_1 \end{cases} \tag{4.47}$$

Then if $\underset{\varepsilon}{\mu} > \frac{p + bp^s}{1 + ap^s} + e\left[a\frac{(p + bp^s)}{(1 + ap^s)} - b\right]$ the expected quality of the new product is greater than its price, and the associated price of the complementary good and the adjustment costs implicit in reducing the capacity of the household technology to cope with a smaller value of the complementary good s .

If $\underset{\varepsilon}{\mu} < \frac{(p + bp^s)}{(1 + ap^s)} - d\left[a\frac{(p + bp^s)}{(1 + ap^s)} - b\right]$, then the expected quality is so low that it is less than the price of the new good and the adjustment cost of increasing the household technology to a level which is implied by consuming only the old good.

If $\frac{p + bp^s}{1 + ap^s} - d\left[a\frac{(p + bp^s)}{(1 + ap^s)} - b\right] < \underset{\varepsilon}{\mu} < \frac{(p + bp^s}{1 + ap^s} + e\left[a\frac{(p + bp^s)}{(1 + ap^s)} - b\right]$

then the consumer maintains his consumption levels of the two goods x and y constant between the two periods, since the expected quality of the new good x , is not great enough to justify the cost of switching to the new good, but neither is the expected quality low enough to justify the costs of returning to the old good.

Then the indirect utility function for the second period can be written:

$$EV_2 = \int_{-\infty}^{\bar{\varepsilon}_1^*} \left\{ \frac{M_2}{1+ap^s} - d\left[a\,\frac{M_2}{1+ap^s} - s_1\right] \right\} f(\bar{\varepsilon})d\bar{\varepsilon}$$

$$+ \int_{\bar{\varepsilon}_1^*}^{\bar{\varepsilon}_2^*} \left[\frac{M_2}{1+ap^s} - \frac{(p+bp^s)}{(1+ap^s)}\,x_1 + \frac{\mu\,x_1}{\bar{\varepsilon}} \right] f(\bar{\varepsilon})\,d\bar{\varepsilon} \qquad (4.48)$$

$$+ \int_{\bar{\varepsilon}_2^*}^{\infty} \left\{ \mu_\varepsilon\,\frac{M_2}{p+bp^s} - e\left[s_1 - \frac{bM_2}{p+bp^s}\right] \right\} f(\bar{\varepsilon})d\bar{\varepsilon}$$

where $\bar{\varepsilon}_1^* = \left\{ \frac{p+bp^s)}{(1+ap^s)} - d\left[a\,\frac{(p+bp^s)}{1+ap^s} - b\right] \right\} \frac{(\phi+x_1)}{x_1} - \frac{\mu_o\phi}{x_1}$

and $\bar{\varepsilon}_2^* = \left\{ \frac{(p+bp^s)}{(1+ap^s} + e\left[a\,\frac{(p+bp^s)}{(1+ap^s)} - b\right] \right\} \frac{\phi+x_1}{x_1} - \frac{\mu_o\phi}{x_1}$

If the consumer starts the first period with a household technology designed for the old good, then s in the previous period was given by $a\,\frac{M_o}{1+ap^s}$. The consumer's problem at the beginning of the first period can be written

$$\max_{s_1,x_1} EU_1 = \frac{M_1}{(1+ap^s)} - \frac{(p+bp^s)}{(1+ap^s)}\,x_1 + \mu_o x_1 - e\left[a\,\frac{M_o}{1+ap^s} - s_1\right] + EV_2$$

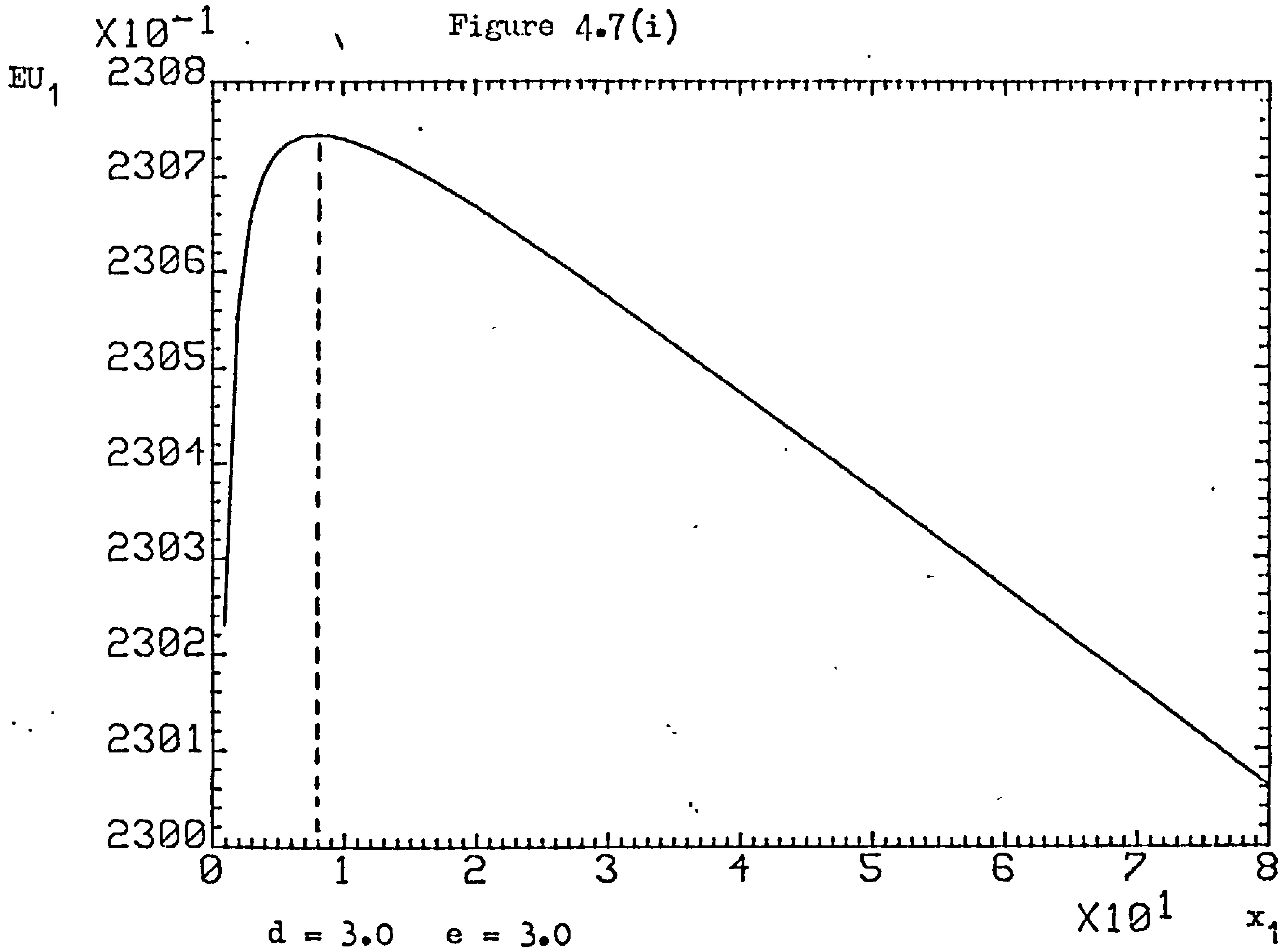
$$\text{s.t.} \quad s_1 \geq ay_1 + bx_1$$

$$\frac{dEU_1}{ds_1} = + e + d \int_{-\infty}^{\bar{\varepsilon}_1^*} f(\bar{\varepsilon})d\bar{\varepsilon}^* - e\int_{\bar{\varepsilon}_2^*} f(\bar{\varepsilon})d\bar{\varepsilon} > 0 \qquad (4.49)$$

So s_1 should be taken to its highest possible value:

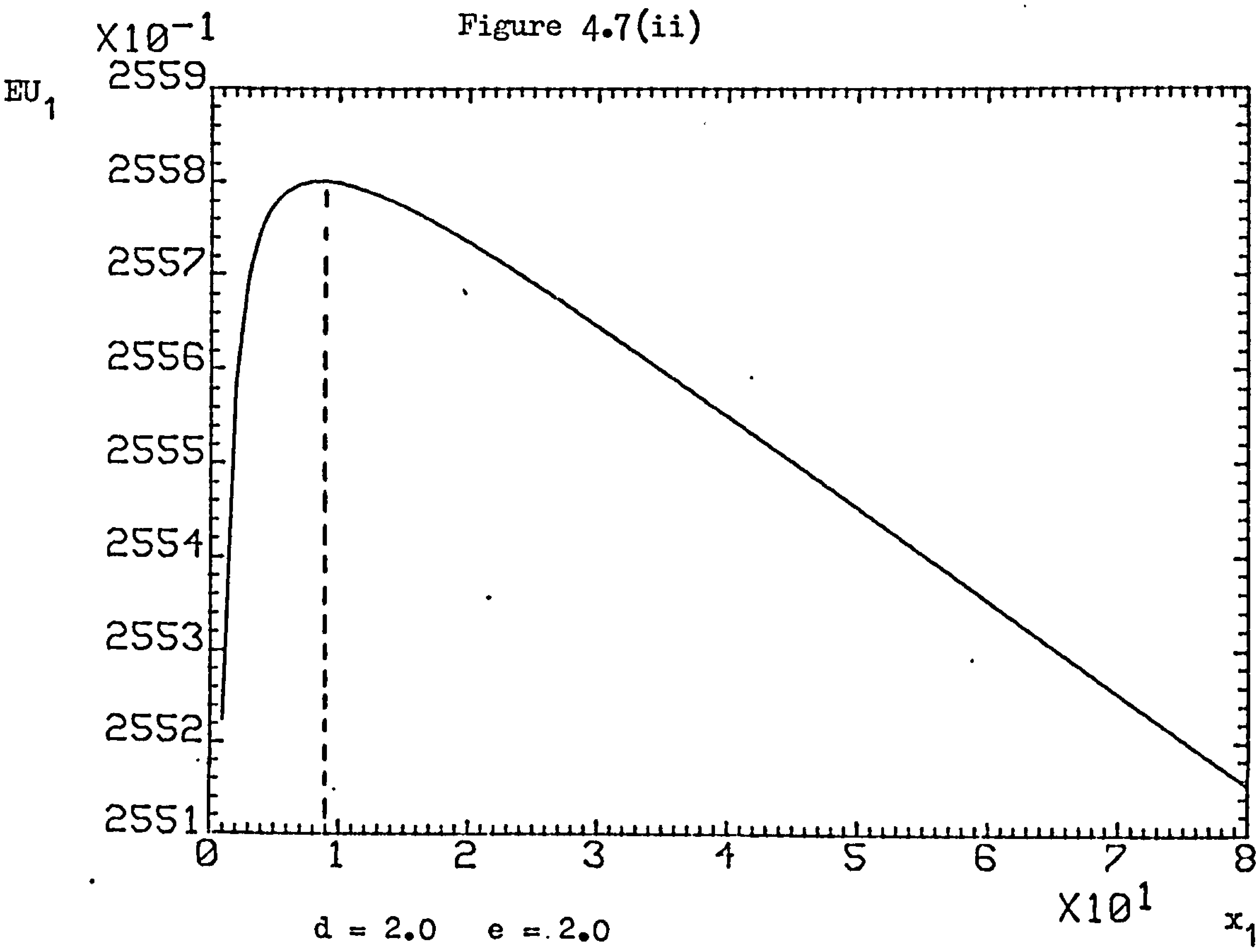
$$s_1 = \frac{a}{1+ap^s}\left[M_1 - (p+bp^s)x_1\right]$$

Figure 4.7(i)



$d = 3.0 \quad e = 3.0$

Figure 4.7(ii)



$d = 2.0 \quad e = 2.0$

Thus the consumer may decide to sample with the new good but will purchase more s_1 than it needs in the first period. Table 5 and Figures 7i and 7ii show a range of parameter values for which the optimal x_1^* is an interior solution. It can be seen that increasing both the adjustment costs reduces the demand for x_1. Since the consumer is purchasing some of the new good, then given the technology relationships in 4.42, the consumer must be

Table 5

μ_o	p	ϕ	M	a	b	p_s	d	e	x_1^*
1.99	6.0	0.01	325	1.5	1.0	2.0	3.0	3.0	8
1.99	6.0	0.01	325	1.5	1.0	2.0	2.0	2.0	9

purchasing more than the necessary minimum quantity of s_1. There exists spare capacity, which is inefficient. It occurs because of the adjustment costs: the consumer is not sure that he will end up purchasing the new good, so therefore he maintains the old level of s in the first period, in case he decides to go back to the old methods in the second.

8. Conclusions

In this chapter we have introduced the way that learning affects behaviour. It was shown that the ability to learn will increase the quantity demanded of the good that the consumer is able to learn about. With the initial model investigated, the decision of the consumer to sample depends upon the parameter values. The introduction of a flexible budget constraint made the existence of the solution less dependent on the specific parameter values.

The consumer's initial beliefs are, in part, provided by an advertising message. Because the consumer recognises that these messages do not furnish accurate information, his beliefs are weighted by the consumer's degree of confidence in the truthfulness of the advertising message, and his experience. The main results in this chapter have been that the opportunity to gain unbiased information, causes the consumer to purchase more of the new good than if he believed with complete confidence, the truthfulness of advertising messages.

We have carried out various comparative exercises on the demand for the new good. Increases in expected quality, price and income all have the expected signs. However the effect of an increase in the degree of confidence is not so clearly seen. Whether an increase in ϕ , causes a rise or fall in x_1^* depends upon the value of the parameters, implying there is some kind of 'optimal' degree of confidence. It will be seen that this result has implications for the multi period analysis in Chapter 7.

We finally looked at the implications of sampling on the efficiency of the household technology, and we showed for some parameter values that the model generated a short-term inefficient solution.

Notes

1. For a more formal definition of a "sufficient" statistic, see DeGroot (1970) p.155.

2. The variance of the random variable ε, could also be unknown; DeGroot (1970) considers the case. But this would only complicate the problem, which is sufficiently rich enough to allow the assumption of a constant variance, with the mean being the only unknown parameter.

3. See Appendix A1.

4. See Appendix A2.

5. Although the mean of ε in equations (4.17) and (4.21) is z, the unconditional distribution is different in each case. The distribution of $\varepsilon \sim N(z, 1)$ is a conditional distribution since it depends upon the value taken by z. To find the unconditional distribution,we need to employ the method used in Appendix A2. This is done in A3.

6. This can be compared with Rothschild (1974) with reference to the Two Armed Bandit Problem

   "If the machine whose pay-off probability is known is ever played, it will be played for evermore", p.191.

7. The concept of a household technology was suggested by Muth (1966).

8. There is a problem in that the $\tilde{\varepsilon}_i$ are discrete observations yet we require x_t to be a continuous variable. According to Cyert, DeGroot and Holt (1978), "we replace the notion of a finite number of discrete projects with the notion of a total investment which is continuously divisible into projects of arbitrary size", p 713. Even if the $\tilde{\varepsilon}_i$ are split into extremely small amounts, each unit will still be a random variable with a known variance.

Appendices

A1 Derivation of posterior mean and variance

See DeGroot (1970) and Theil (1978).

$\varepsilon_1 \ldots\ldots \varepsilon_x$ is a random sample, where each of the ε_i are independently and identically distributed random variables with unknown mean x and unit variance, $f(\varepsilon|z,1)$. The prior distribution of z is normal with mean μ_o and precision ϕ, $f(z|\mu_o, \frac{1}{\phi})$.

Bayes theorem states:

$$f(z|\varepsilon_1, \ldots \varepsilon_x) = \frac{f(z|\mu_o,\frac{1}{\phi})\, L(\varepsilon_1, \ldots \varepsilon_x|z)}{\int_{\forall z} L(\varepsilon_1, \ldots \varepsilon_x|z)\, f(z|\mu_o,\frac{1}{\phi})dz}$$

where $L(\varepsilon_1, \ldots\ldots \varepsilon_x|z)$ is the likelihood function for the random sample $\varepsilon_1, \ldots\ldots \varepsilon_x$. It is the joint density function of the random variable from a normal distribution with mean z and unit variance. It is the probability density of obtaining the sample given the value of the parameter z. We may write out the likelihood function:

$$\begin{aligned} L(\varepsilon_1 \ldots\ldots \varepsilon_x|z) &= f(\varepsilon_1|z,1) \cdot f(\varepsilon_2|z,1) \ldots\ldots f(\varepsilon_x|z,1) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\varepsilon_1 - z)^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\varepsilon_2 - z)^2} \ldots\ldots \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\varepsilon_x - x)^2} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^x \exp\left\{-\frac{1}{2} \sum_{i=1}^{x} (\varepsilon_i - z)^2\right\} \end{aligned}$$

and $$f(z|\mu_o,\tfrac{1}{\phi}) = \sqrt{\frac{\phi}{2\pi}} \exp\left\{ -\frac{\phi}{2} (z - \mu_o)^2 \right\}$$

We may write Bayes' theorem as:

$$f(z|\varepsilon_1 \dots \varepsilon_x) \propto \sqrt{\frac{\phi}{2\pi}} \exp\{-\frac{\phi}{2}(z-\mu_o)^2 \cdot \frac{1}{(\sqrt{2\pi})^x} \exp\{-\frac{1}{2} \sum_{i=1}^{x} (\varepsilon_i - z)^2\}$$

Taking logs, the proportionality constant represented by $\propto$ becomes an additive constant, with respect to the unknown parameter z.

$$\therefore \log_e f(z|\varepsilon_1 \dots \varepsilon_x) = -\frac{\phi}{2}(z-\mu_o)^2 - \frac{1}{2} \sum_{i=1}^{x} (\varepsilon_i - z)^2 + \text{constant}$$

and
$$\sum_{i=1}^{x} (\varepsilon_i - z)^2 = \sum_{i=1}^{x} \left[(\varepsilon_i - \bar{\varepsilon}) + (\bar{\varepsilon} - z)\right]^2$$

$$= \sum_{i=1}^{x} (\varepsilon_i - \bar{\varepsilon})^2 + x(\bar{\varepsilon} - z)^2 + 2(\bar{\varepsilon} - z) \sum_{i=1}^{x} (\varepsilon_i - \bar{\varepsilon})$$

But $\sum_{i=1}^{x} (\varepsilon_i - \bar{\varepsilon}) = 0$, and $\sum_{i=1}^{x} (\varepsilon_i - \bar{\varepsilon})^2$ is a constant with respect to z; so,

$$\log_e f(z|\varepsilon_1 \dots \varepsilon_x) = -\frac{1}{2}\{\phi(z-\mu_o)^2 + x(\bar{\varepsilon} - z)^2\} + \text{constant}$$

And
$$\phi(z-\mu_o)^2 + x(\bar{\varepsilon} - z)^2 = \phi(z^2 - 2\mu_o z + \mu_o^2) + (\bar{\varepsilon}^2 - 2\bar{\varepsilon}z + z^2)$$

$$= (\phi + x)\left(z^2 - 2z\frac{(\phi\mu_o + x\bar{\varepsilon})}{\phi + x} + \frac{\phi\mu_o^2 + x\bar{\varepsilon}^2}{\phi + x}\right)$$

Completing the square
$$= (\phi + x)\left[z - \frac{(\phi\mu_o + \bar{\varepsilon}x)}{\phi + x}\right]^2 + (\phi + x)\left[\frac{\phi\mu_o^2 + x\bar{\varepsilon}^2}{\phi + x} - \left(\frac{\phi\mu_o + x\bar{\varepsilon}}{\phi + x}\right)^2\right]$$

But the final expression is again constant w.r.t. z.

$$\therefore \log_e f(z|\varepsilon_1 \ldots \varepsilon_x) = -\frac{1}{2}\{(\phi+x)\left[z - \frac{(\phi\mu_o+\bar{\varepsilon}x)}{\phi+x}\right]^2\} + \text{constant}$$

Taking anti-logs:

$$f(z|\varepsilon_1 \ldots \varepsilon_x) \propto \exp\{-\frac{1}{2}(\phi+x)\left[z - \frac{\phi\mu_o+\bar{\varepsilon}x}{\phi+x}\right]^2\}$$

which is of the form of a normal distribution with the mean $\mu_{\bar{\varepsilon}}$ and variance $\sigma^2_{\bar{\varepsilon}}$ specified in equation (4.15) and (4.16).

A2 <u>Derivation of marginal distribution of $\bar{\varepsilon}$.</u>

Conditional distribution of $\bar{\varepsilon}$ is $N(z, \frac{1}{x})$, with density function $f_{\bar{\varepsilon}|z}$, which can be derived from equation (4.14) since $\varepsilon_i \sim N(z, 1)$.

Prior distribution of z is $N(\mu_o, \frac{1}{\phi})$ with density function f_z.

If joint distribution of z and $\bar{\varepsilon}$ is written $f_{\bar{\varepsilon},z}$, then

$$f_{\bar{\varepsilon}|z} \cdot f_z = f_{\bar{\varepsilon},z}$$

and marginal distribution of $\bar{\varepsilon}$ is

$$f_{\bar{\varepsilon}} = \int_{\forall z} f_{\bar{\varepsilon},z}\, dz$$

$$\therefore f_{\bar{\varepsilon}} = \int_{-\infty}^{\infty} \sqrt{\frac{x}{2\pi}} \exp\{-\frac{x}{2}(\bar{\varepsilon}-z)^2\} \cdot \sqrt{\frac{\phi}{2\pi}} \exp\{-\frac{\phi}{2}(z-\mu_o)^2\}\, dz$$

$$= \frac{\sqrt{\phi x}}{2\pi} \int_{-\infty}^{\infty} \exp -\frac{1}{2}\{x(\bar{\varepsilon}-z)^2 + (x-\mu_o)^2\}\, dz$$

$$= \frac{\sqrt{\phi x}}{2\pi} \int_{-\infty}^{\infty} \exp -\frac{1}{2}\{(x+\phi)\left[z^2 - 2z\frac{(\phi\mu_o+\bar{\varepsilon}x)}{\phi+x} + \frac{\phi\mu_o^2+\bar{\varepsilon}^2 x}{\phi+x}\right]\} dz$$

Completing the square

$$= \sqrt{\frac{\phi x}{2\pi}} \int_{-\infty}^{\infty} \exp - \frac{1}{2} \{(\phi+x)\left[z - \frac{\phi\mu_o + \bar{\varepsilon}x}{\phi+x}\right]^2 + \phi\mu_o^{\,2} + x\bar{\varepsilon}^2 + \frac{(\phi\mu_o+\bar{\varepsilon}x)^2}{\phi + x}\} \, dz$$

$$= \sqrt{\frac{\phi+x}{2\pi}} \int_{-\infty}^{\infty} \exp - \frac{(\phi+x)}{2}\left[z - \frac{\phi\mu_o+\bar{\varepsilon}x}{\phi + x}\right]^2 dz \, . \sqrt{\frac{1}{2\pi} \cdot \frac{\phi x}{\phi+x}} \cdot \exp - \frac{1}{2} \frac{\phi x}{\phi+x}(\bar{\varepsilon}-\mu_o)^2$$

$$\therefore \; f_{\bar{\varepsilon}} = \sqrt{\frac{1}{2\pi} \cdot \frac{\phi x}{\phi+x}} \cdot \exp - \frac{1}{2} \{ \frac{\phi x}{\phi+x} \cdot (\bar{\varepsilon} - \mu_o)^2\}$$

which is a normal distribution with mean μ_o and variance $\frac{1}{\phi} + \frac{1}{x}$.

since $\sqrt{\frac{\phi+x}{2\pi}} \int_{-\infty}^{\infty} \exp - \frac{1}{2} \{(\phi+x)\left[z - \frac{\phi\mu_o + \bar{\varepsilon}x}{\phi + x}\right]^2$ is the area under a normal distribution with mean $\frac{\phi\mu_o + \bar{\varepsilon}x}{\phi + x}$ and variance $\frac{1}{\phi + x}$.

A3 <u>Derivation of the marginal distribution of ε.</u>

Conditional distribution of ε is $N\,(z,1)$ and prior distribution of z is $N\,(\mu_o,\frac{1}{\phi})$. Then in the first period, marginal distribution of ε is

$$f^1_\varepsilon = \int_{\forall z} f\varepsilon|z \;\; f^1_z \;\; dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}(\varepsilon - z)^2\} \, . \sqrt{\frac{\phi}{2\pi}} \exp \; -\frac{\phi}{2}(z - \mu_o)^2 \quad dz$$

Using the same method employed in A2

$$f^1_\varepsilon = \sqrt{\frac{1}{2\pi} \; \frac{\phi}{(\phi+1)}} \cdot \exp - \frac{1}{2} \{ \frac{\phi}{(\phi+1)} \;\; (\varepsilon - \mu_o)^2 \}$$

which is a normal distribution with mean μ_o and variance $\frac{1+\phi}{\phi}$.

In the second period, marginal distribution of ε is

$$f_\varepsilon^2 = \int_{\forall z} f\varepsilon|z \; f_z^2 \, dz$$

where f_z^2 is the posterior distribution of z.

$$f^2 = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp - \frac{1}{2}(\varepsilon - z)^2 . \sqrt{\frac{\phi + x}{2\pi}} \exp - \frac{1}{2}(\phi + x)(z - \mu_\varepsilon)^2 \, dz$$

Again using the method employed in A2

$$f_\varepsilon^2 = \sqrt{\frac{1}{2\pi} \cdot \frac{\phi + x}{1 + \phi + x}} \exp - \frac{1}{2} \left\{ \frac{(\phi + x)}{\phi + x + 1} (\varepsilon - \mu_o)^2 \right\}$$

which is again a normal distribution with mean μ_o and variance $\frac{\phi + x + 1}{\phi + x}$.

A4 Proof to Proposition 4.2

Differentiate equation (4.22) w.r.t. x_1, using (4.25) and noting

$$\frac{dm^*}{dx_1} = -(p - \mu_o)\left(\frac{\phi}{x_1}\right)^{3/2} \cdot \frac{1}{(\phi + x_1)^{\frac{1}{2}}}$$

and $\int_{m^*}^{\infty} mf(m|0,1) \, dm = f(m^*)$

Then $$\frac{d^2 EU_1}{dx_1^2} = \frac{M_2}{4p} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \cdot \frac{f(m^*)}{x_1} \cdot \frac{1}{(\phi + x_1)^{3/2}} \left[(p - \mu_o)^2 \frac{\phi}{x_1}^2 - \frac{(\phi + 4x_1)}{(\phi + x_1)} \right]$$

which has a negative and a positive component. When $x_1 = 0$, the positive term approaches infinity, so that at low values of x_1, EU_1 is concave. As x_1 increases the function becomes convex.

Q. E. D.

A5 <u>Proof to Proposition 4.3</u>

We know from proposition 1 that $\theta > 0$.

(iv) If $\mu_o \geq p$ then $\theta + \mu_o > p$ so $\frac{dEU_1}{dx_1} > 0 \; \forall \; x_1$.

For (i) - (iii) $\mu_o < p$ and the sign of $\frac{dEU_1}{dx_1}$ will depend upon the absolute size of θ, the information component, i.e. will the positive θ be sufficient to outweigh the negative $(\mu_o - p)$

(i) If θ is small such that $\mu_o + \theta < p$ then $\frac{dEU_1}{dx_1} < 0 \; \forall \; x_1$

(iii) If θ is large such that $\mu_o + \theta > p$ then $\frac{dEU_1}{dx_1} > 0$, but θ depends upon the value of x_1. So only at high values of x_1: $\frac{dEU_1}{dx_1} > 0$.

When $x_1 = 0 \rightarrow \theta = 0$ and so for small values of x_1, $\frac{dEU_1}{dx_1} < 0$.

(ii) But at very high values of x_1, $\theta = 0$, as can be seen from equation (4.24).

In this case $x_1 = 0 \rightarrow \frac{dEU_1}{dx_1} < 0$

and $x_1 = \frac{M_1}{p} \rightarrow \frac{dEU_1}{dx_1} < 0$,

however there will be intermediate values of x_1 for which $\frac{dEU_1}{dx_1} > 0$.

Whichever effect dominates depends upon the parameter values.

<u>Q. E. D.</u>

A6 Proof to Proposition 4.4

Differentiate equation (4.21) w.r.t. ϕ, where EV_2 is given by (4.24)

$$\frac{dEU_1}{d\phi} = \frac{-1}{1+i} \cdot \frac{M_2}{p} \cdot \frac{(x_1 + 2\phi)}{(x_1 + \phi)^{3/2}} \cdot \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{f(m^*)}{\phi} < 0$$

Q. E. D.

A7 Proof to Proposition 4.5

Implicitly differentiate (4.22) using (4.25)

$$\frac{dx_1}{d\phi} = -\frac{1}{\frac{d^2EU_1}{dx_1^{\,2}}} \left\{ \frac{1}{1+i} \cdot \frac{d^2EV_2}{dx_1 d\phi} \right\}$$

We assume that the second order conditions for an optimum x_1^* are satisfied so that $\frac{d^2EU_1}{dx_1^{\,2}} < 0$.

Then $$\frac{d^2EV_2}{dx_1 d\phi} = \frac{M_2}{4p} \cdot \frac{f(m^*)}{(x_1+\phi)^{3/2}} \cdot \frac{1}{(\phi x_1)^{\frac{1}{2}}} \left\{ \frac{x_1-2\phi}{x_1+\phi} - (p-\mu_o)^2 \cdot \frac{\phi}{x_1} \cdot (x_1+2\phi) \right\}$$

So $\frac{dx_1}{d\phi} \gtrless 0$ as $x_1^{\,2} \gtrless (p-\mu_o)^2 \, \phi \, (x_1^2 + 3\phi x_1 + 2\phi^2) + 2\phi x_1$

A8 Proof to Proposition 4.6

Implicitly differentiate equation (4.22)

$$\frac{dx_1}{d\mu_o} = -\frac{1}{\frac{d^2EU_1}{dx_1^{\,2}}} \left\{ 1 + \frac{1}{1+i} \cdot \frac{dEV_2^{\,2}}{dx_1 d\mu_o} \right\}$$

Differentiation (4.25) w.r.t. μ_o

$$\frac{dEV_2^{\,2}}{dx_1 d\mu_o} = \frac{M_2}{2p} \cdot \frac{(p - \mu_o)}{(\phi + x_1)} \cdot \left(\frac{\phi}{x_1}\right)^{3/2} f(m^*) > 0$$

Q. E. D.

A9 Proof to Proposition 4.7

Implicitly differentiate equation (4.22)

$$\frac{dx_1}{dM_2} = \frac{-1}{\dfrac{d^2EU_1}{dX_2^{\,2}}} \left\{ \frac{1}{1+i} \frac{dEV_2^{\,2}}{dM_2 dx_1} \right\}$$

Differentiating (4.25) w.r.t. M_2

$$\frac{dEV_2^{\,2}}{dx_1 dM_2} = \frac{1}{2p} \frac{1}{(\phi+x_1)^{3/2}} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \cdot f(m^*) > 0$$

Q. E. D.

A10 Proof to Proposition 4.8

Implicitly differentiate equation (4.22)

$$\frac{dx_1}{dp} = \frac{-1}{\dfrac{d^2EU_1}{dx_2^{\,2}}} \left\{ -1 + \frac{1}{1+i} \frac{d^2EV_2}{dx_1 dp} \right\}$$

Differentiate (4.25) w.r.t. p

$$\frac{d^2EV_2}{dx_1 dp} = \frac{M_2}{2} \frac{1}{(\phi+x_1)^{3/2}} \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \left[-\frac{f(m^*)}{p^2} - f(m^*) \frac{m^*}{p} \frac{dm^*}{dp} \right]$$

$$= -\frac{M_2}{2p} \cdot \frac{f(m^*)}{(\phi+x_1)^{3/2}} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \cdot \left[\frac{1}{p} + (p - \mu_o)(\phi+x_1) \frac{\phi}{x_1} \right]$$

$\therefore \dfrac{d^2EV_2}{dx_1 dp} < 0$ since $p > \mu_o$ for internal solution.

Q. E. D.

A11 Proof to Proposition 4.9

Carrying out the maximisation procedure in (4.30), first order conditions yield:

$$\frac{dEU_1'}{dy_1} = 1 + \frac{1}{1+i} \cdot \frac{dEV_2'}{dy_1}$$

Differentiate (4.32) w.r.t. y_1 :

$$\frac{dEV_2'}{dy_1} = (1+i)\left[- \int_{-\infty}^{m^*} f(m)dm - \frac{1}{p} \int_{m^*}^{\infty} \left[\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right] f(m)dm \right]$$

Making the substitution $\int_{-\infty}^{m^*} f(m)dm = 1 - \int_{m^*}^{\infty} f(m)dm$

Then $$\frac{dEU_1'}{dy_1} = \int_{m^*}^{\infty} \{ 1 - \frac{1}{p} \left[\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right] \} f(m)dm$$

But for values of m in the range (m^*, ∞)

$$p < \mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}}$$

So

$$\frac{dEU_1'}{dy_1} < 0 \rightarrow y_1^* = 0$$

Q. E. D.

A12 Proof to Proposition 4.10

Maximise equation (4.30) w.r.t. x_1,

$$\frac{dEU_1'}{dx_1} = \mu_o + \frac{1}{1+i} \frac{dEV_2'}{dx_1}$$

Differentiate (4.32) w.r.t. x_1,

$$\frac{dEV_2'}{dx_1} = \frac{(1+i)(M-px_1)}{2p} \cdot \frac{1}{(\phi+x_1)^{3/2}} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \int_{m^*}^{\infty} mf(m)dm.$$

$$- (1+i)\left[\int_{m^*}^{\infty} \{\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\} f(m)dm + p \int_{-\infty}^{m^*} f(m)dm\right]$$

But $$\int_{-\infty}^{+\infty} \left[\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right] f(m)dm = \mu_o$$

$$\therefore \frac{dEU_1'}{dx_1} = \frac{(M-px_1)}{2p} \frac{f(m^*)}{(\phi+x_1)^{3/2}} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} + \int_{-\infty}^{m^*} \left[\mu_o - p + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right] f(m)dm$$

But for values of m in the range $(-\infty, m^*)$

$$p > \mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x)^{\frac{1}{2}}}$$

So, at $x_1 = \frac{M}{p}$

$$\frac{dEU_1'}{dx_1} = 0 + \int_{-\infty}^{m^*} \left[\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}} - p\right] f(m)dm < 0$$

$$\therefore x_1^* < \frac{M}{p}$$

Q. E. D.

A13 Proof to Proposition 4.11

Differentiate (4.38) using (4.39) and (4.40) w.r.t. x_1

$$\frac{dEU_1''}{dx_1} = \mu_o + \frac{(M-p_1x_1)}{2p_2} \cdot \frac{1}{(\phi+x_1)^{3/2}} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \int_{m^*}^{\infty} mf(m)dm$$

$$- \frac{p_1}{p_2} \int_{m^*}^{\infty} \left[\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right] f(m)dm - p_1 \int_{-\infty}^{m^*} f(m)dm$$

$$= \left[1 - \frac{p_1}{p_2}\right] \mu_o + \frac{(M-px_1)}{2p_2} \cdot \frac{1}{(\phi+x_1)^{3/2}} \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \int_{m^*}^{\infty} mf(m)dm$$

$$+ \frac{p_1}{p_2} \int_{-\infty}^{m^*} \left[\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}} - p_2\right] f(m)dm$$

As $m^* \to \infty$: $\displaystyle\int_{-\infty}^{m^*} \left[\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \frac{m}{(\phi+x_1)^{\frac{1}{2}}}\right] f(m)dm \to \mu_o$

Then $\displaystyle\frac{dEU_1''}{dx_1} = \mu_o - p_1 > 0$

So $\displaystyle x_1^* = \frac{M}{p_1}$

Q. E. D.

A14 Proof to Proposition 4.12

Implicitly differentiate the F.O.C.'s in A13.

$$\frac{dx_1}{dp_2} = - \frac{1}{\frac{d^2EU_1''}{dx_1{}^2}} \cdot \frac{d^2EV_2''}{dx_1dp_2}$$

$$\frac{d^2EV_2''}{dx_1dp_2} = \frac{-(M-p_1x_1)}{2p_2} \cdot \frac{1}{(\phi+x_1)^{3/2}} \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} m^*f(m^*) \frac{dm^*}{dp_2}$$

$$- \frac{M-p_1x_1}{2p_2^{\;2}} \cdot \frac{f(m^*)}{(\phi+x_1)^{3/2}} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}}$$

$$+ \frac{p_1}{p_2} \int_{m^*}^{\infty} \left[\mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi+x_1)^{\frac{1}{2}}} \right] f(m)dm$$

Divide F.O.C.'s in A13 by p_2, and substitute for

$$\frac{M-p_1x_1}{2p_2^{\;2}} \cdot \frac{f(m^*)}{(\phi+x_1)^{3/2}} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}}$$

Then $$\frac{d^2EV_2''}{dx_1dp_2} = \frac{(M-p_1x_1)}{2p_2} \cdot \frac{p_2-\mu_o}{(\phi+x_1)^{\frac{1}{2}}} \cdot \left(\frac{\phi}{x_1}\right)^{3/2} f(m^*) + \frac{\mu_o}{p_2} - \frac{p_1}{p_2} \int_{-\infty}^{m^*} f(m)dm$$

Let $$\int_{-\infty}^{m^*} f(m)dm = F(m^*)$$

Then $$\frac{d^2EV_2''}{dx_1dp_2} \gtrless 0 \text{ as } \mu_o\left[1+\frac{(M-p_1x_1)}{2} \cdot \frac{f(m^*)}{(\phi+x_1)^{\frac{1}{2}}} \cdot \left(\frac{\phi}{x_1}\right)^{3/2}\right]$$

$$\gtrless \left[p_1F(m^*) + \frac{p_2(M-p_1x_1)}{(\phi+x_1)^{\frac{1}{2}}} \cdot \left(\frac{\phi}{x_1}\right)^{3/2} f(m^*) \right]$$

Q. E. D.

Chapter 5

Learning with a C.E.S. Utility Function

1. Introduction

In this chapter we follow the methodology of the previous chapter in examining the effect of learning on consumer behaviour. But we change the circumstances of the consumer in two ways

a) The utility function is specified as belonging to the class of constant elasticities of substitution.

b) The random variable is assumed to have a log-normal distribution.

These two differences introduce important characteristics to the problem. Firstly a C.E.S. utility function is strictly concave, unlike the previously considered linear one, which has implications for the risk attitudes of the consumer. Secondly, the log-normal distribution is not symmetric around its mean; so that the parameter supplied by the firm μ, is not used as the center of the distribution.

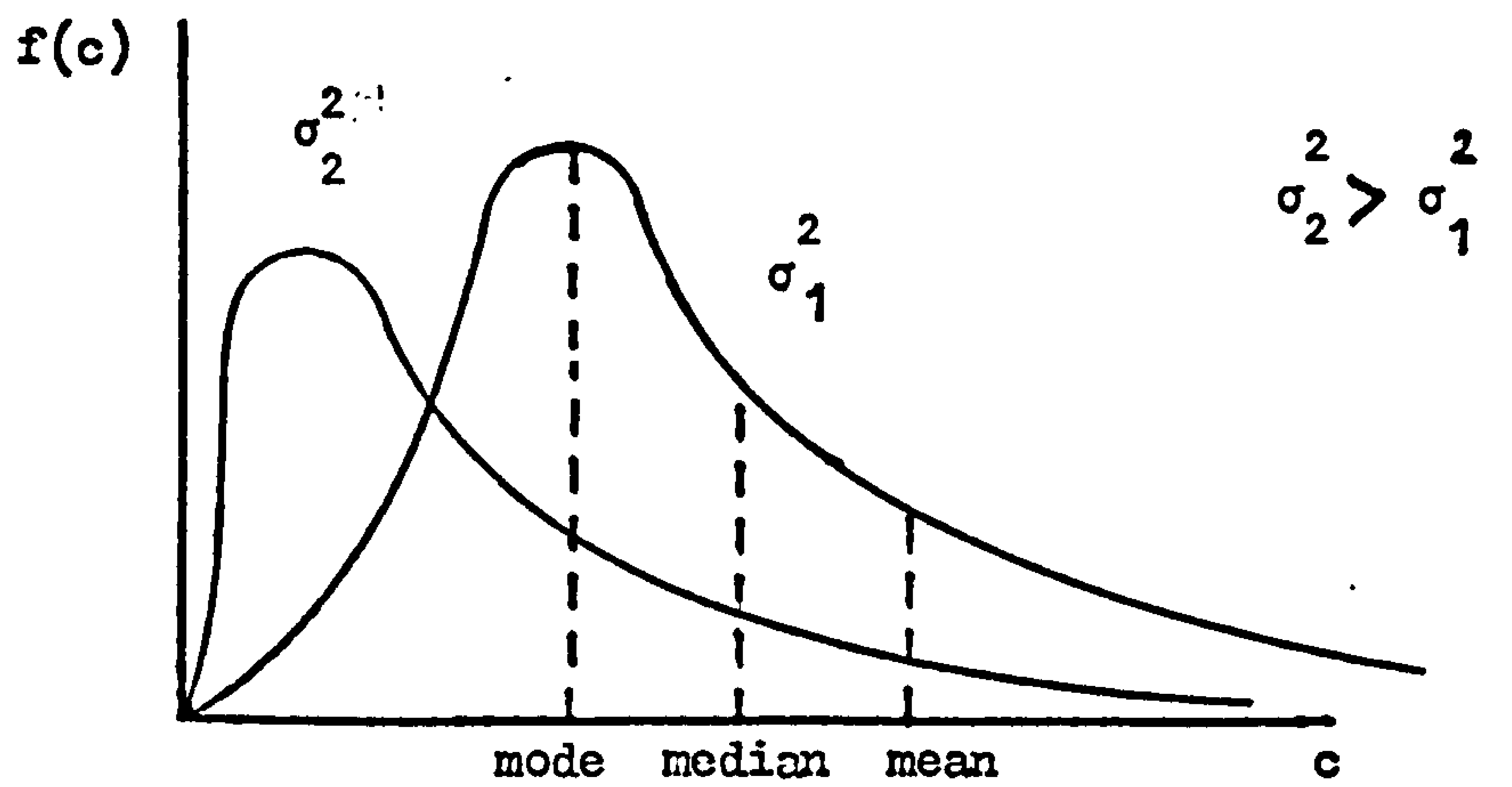
We can then ask the question: do the results of the last chapter alter as a consequence of these changes?

We would not expect the introduction of risk aversion to greatly affect the main results from Chapter 4. In fact Grossman et al (1977) considered a general risk-averse utility function in proving their Theorem 1. It is worthwhile emphasising this result in the context of a concave utility function. The implication of learning on the size of the decision variable in the first period is not that the consumer will always choose to purchase a "risky" new good. Whether the consumer buys the new good or not depends upon his attitudes towards risk.

What the result does say, is that if the consumer is able to learn about the uncertainty, he will purchase more than if that opportunity for learning does not exist.

Consider a simple portfolio choice problem: the investor must choose between a safe asset and a risky asset. Suppose firstly that the distribution of the returns from the risky asset are known; then the investor's preferences over risk and return will determine his holdings of the two assets. Compare this with the situation when the distribution of returns on the risky asset is adaptive. Then Grossman et al and our results in Chapter 4 say that the investor will purchase more of the risky asset in the second case than the first.

The general properties of the log-normal distribution have been extensively discussed in Aitchison and Brown (1957). If $c = e^{-\varepsilon}$, such that $\log_e c = -\varepsilon$ and ε is a random variable with a normal distribution $N(\mu, \sigma^2)$ then c has a lognormal distribution and parameters μ and σ^2. The mean of this distribution is $e^{-\mu + \frac{\sigma^2}{2}}$, and c takes on only positive values $(0 < c < \infty)$.



mode $= e^{-\mu - \sigma^2}$

median $= e^{-\mu}$

mean $= e^{-\mu + \frac{\sigma^2}{2}}$

The distribution is skewed; its mean lies to the right of both the mode and median. Thus the consumer knows that most observations from this distribution will lie to the left of the mean. We still suppose that the consumer's beliefs about the truth of the advertising message, are normally distributed, but the underlying distribution which the advertising message is purporting to inform the consumer about is skewed. We might expect that the importance of this fact, is that it is more difficult for a consumer to verify the truth of an advertising statement, since he expects most of the observations to be concentrated below the mean.

2. The Model

Suppose that in each period the consumer faces the following utility function:

$$u_t = -\frac{\alpha}{y_t} - \frac{\beta e^{-\varepsilon}}{x_t} \qquad (5.1)$$

$\varepsilon \sim N(z, {}^1/_r)$ but z is unknown with distribution $\sim N(\mu_0, {}^1/_\phi)$

$$\therefore \quad EU_t = -\frac{\alpha}{y_t} - \frac{\beta e^{-z+\frac{1}{2r}}}{x_t}$$

In the first period

$$EU_1 = -\frac{\alpha}{y_1} - \frac{\beta e^{-\mu_0+\frac{1}{2\phi}+\frac{1}{2r}}}{x_1}$$

But in the second period $z \sim N(\mu_1, \sigma_1^2)$

where $\mu_1 = \dfrac{\mu_0\phi + \bar{\varepsilon}x_1}{\phi + x_1}$ and $\sigma_1^2 = \dfrac{1}{\phi + x_1}$

So $$EU_2 = -\frac{\alpha}{y_2} - \frac{\beta e^{-\mu_1+\frac{1}{2}\sigma_1^2+\frac{1}{2r}}}{x_2}$$

$$= -\frac{\alpha}{y_2} - \frac{\beta e^{-\theta}}{x_2} \quad \text{where } \theta = \mu_1 - \frac{1}{2}\sigma_1^2 - \frac{1}{2r}$$

$$\max \quad EU_2 = -\frac{\alpha}{y_2} - \frac{\beta e^{-\theta}}{x_2} + \lambda(M_2 - p_y y_2 - p_x x_2) \tag{5.2}$$

First order conditions yield:

$$\therefore \frac{dEU_2}{dy_2} = \frac{\alpha}{y_2^2} - \lambda p_y = 0$$

$$\frac{dEU_2}{dx_1} = \frac{\beta e^{-\theta}}{x_2^2} - \lambda p_x = 0$$

$$\frac{\alpha}{\beta} \quad \frac{x_2^2 e^{\theta}}{y_2^2} = \frac{p_y}{p_x}$$

$$\therefore x_2 = y_2 \sqrt{\frac{p_y}{p_x} \cdot \frac{\beta}{\alpha} e^{-\theta}} \tag{5.3}$$

Substitute into budget constraint $M = p_y y + p_x x$

$$\therefore y_2 = \frac{M}{p_y} - \frac{p_x}{p_y} \cdot y_2 \sqrt{\frac{p_y}{p_x} \frac{\beta e^{-\theta}}{\alpha}}$$

$$\therefore y_2 = \frac{M}{p_y + \sqrt{p_x p_y \frac{\beta}{\alpha} e^{-\theta}}} \tag{5.4}$$

Substitute for y_2 in (5.3)

$$x_2 = \frac{M}{p_x + (p_y \; p_x \frac{\alpha}{\beta})^{1/2} e^{\theta/2}} \tag{5.5}$$

Substitute back into equation (5.2)

$$V_2 = -\frac{\alpha}{M}\left[p_y + (p_x p_y \frac{\beta}{\alpha})^{1/2} e^{-\theta/2}\right] - \frac{\beta}{M}\left[p_x + (p_y p_x \frac{\alpha}{\beta})^{1/2} e^{\theta/2}\right] e^{-\theta}$$

$$= -\frac{1}{M}\left\{\alpha p_y + (\alpha\beta \; p_x p_y)^{1/2} e^{-\theta/2} + \beta p_x e^{-\theta} + (\alpha\beta p_x p_y)^{1/2} e^{-\theta/2}\right\}$$

$$= -\frac{1}{M}\left\{\alpha p_y + \beta p_x e^{-\theta} + 2(\alpha\beta \; p_x p_y)^{1/2} e^{-\theta/2}\right\} \tag{5.6}$$

But θ depends upon the random variable $\bar{\varepsilon}$, thus to find the expected value of the indirect utility function in period 2 we need to integrate over the random variable.

$$EV_2 = -\frac{1}{M}\int_{-\infty}^{\infty} (\alpha p_y + \beta p_x e^{-\theta} + 2(\alpha\beta p_x p_y)^{1/2} e^{-\theta/2})\, f(\bar{\varepsilon}|\mu_0, \frac{1}{\phi} + \frac{1}{rx_1})\, d\bar{\varepsilon} \quad (5.6$$

Following Theil (1971)

$$\int_{-\infty}^{+\infty} \bar{\varepsilon}^{\,t\bar{\varepsilon}} f(\bar{\varepsilon}|\mu_0, \sigma^2)\, d\bar{\varepsilon} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left\{-t\bar{\varepsilon} - \frac{1}{2}\frac{(\bar{\varepsilon}-\mu_0)^2}{\sigma^2}\right\} d\bar{\varepsilon}$$

But

$$-t\bar{\varepsilon} - \frac{1}{2}\frac{(\bar{\varepsilon}-\mu_0)^2}{\sigma^2} = -\frac{1}{2}\cdot\frac{(\bar{\varepsilon}-\mu_0)^2 + 2\sigma^2 t\bar{\varepsilon}}{\sigma^2}$$

$$= -\frac{1}{2}\cdot\frac{(\bar{\varepsilon}-\mu_0+\sigma^2 t)^2 + 2\mu_0\sigma^2 t - \sigma^4 t^2}{\sigma^2}$$

$$= -\frac{1}{2}\cdot\frac{(\bar{\varepsilon}-\mu_0+\sigma^2 t)^2}{\sigma^2} - \mu_0 t + \frac{\sigma^2 t^2}{2}$$

Thus $\bar{\varepsilon}$ now becomes a random variable with mean $\mu_0 - \sigma^2 t$ and variance σ^2, and the area under a normal curve sums to 1.

$$\therefore \int_{-\infty}^{+\infty} e^{-t\bar{\varepsilon}} f(\bar{\varepsilon}|\mu_0, \frac{1}{\phi} + \frac{1}{rx_1})\, d\bar{\varepsilon} = \exp\left\{-\mu_0 t + \frac{(rx_1+\phi)}{2\phi\, x_1 r} t^2\right\}$$

Applying this result to the integration problem in (5.6):

$$\int_{-\infty}^{+\infty} \exp\left\{-\frac{(\mu_0\phi + \bar{\varepsilon}\, x_1 r)}{rx_1+\phi} + \frac{1}{2(rx_1+\phi)} + \frac{1}{2r}\right\} f(\bar{\varepsilon}|\mu_0, \frac{1}{\phi} + \frac{1}{rx_1})\, d\bar{\varepsilon}$$

$$= \exp\left\{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}\right\}$$

and

$$\int_{-\infty}^{+\infty} \exp\left\{-\frac{(\mu_0\phi + \bar{\varepsilon}\, x_1 r)}{2(rx_1+\phi)} + \frac{1}{4(rx_1+\phi)} + \frac{1}{4r}\right\} f(\bar{\varepsilon}|\,\cdot)\, d\bar{\varepsilon} =$$

$$\exp\left\{-\frac{\mu_0}{2} + \frac{(\frac{1}{2}x_1 r + \phi)}{4\phi(rx_1+\phi)} + \frac{1}{4r}\right\}$$

Equation (5.6) can be written:

$$EV_2 = -\frac{1}{M}\left\{\alpha p_y + \beta p_x e^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}} + 2(\alpha\beta p_x p_y)^{1/2} e^{-\frac{\mu_0}{2} + v + \frac{1}{4r}}\right\} \quad (5.7)$$

where

$$v = \frac{1}{4}\cdot\frac{1/2\, rx_1 + \phi}{\phi(rx_1+\phi)}$$

Writing out the consumer's two period objective:

$$\max_{x_1, y_1} EU_1 = -\frac{\alpha}{y_1} - \frac{\beta e}{x_1}^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}} + EV_2(x_1) \tag{5.8}$$

$$\text{s.t.} \quad M_1 = p_y y_1 + p_x x_1$$

<u>**F.O.C.**</u>

$$\left.\begin{aligned} \therefore \quad & \frac{dEU_1}{dy_1} = \frac{\alpha}{y_1^2} - \lambda p_y = 0 \\ & \frac{dEU_1}{dx_1} = \frac{\beta e}{x_1^2}^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}} + \frac{dEV_2}{dx_1} - \lambda p_x = 0 \end{aligned}\right\} \tag{5.9}$$

$$\therefore \quad \frac{\alpha}{y_1^2\left[\frac{\beta e}{x_1^2}^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}} + \frac{dEV_2}{dx_1}\right]} = \frac{p_y}{p_x} \tag{5.10}$$

$$\therefore \quad y_1 = \left\{\frac{\alpha\, p_x}{p_y\left[\frac{\beta e}{x_1^2}^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}} + \frac{dEV_2}{dx_1}\right]}\right\}^{1/2}$$

But from $(5.9)_1$ α/y_1^2 is the marginal utility of one unit of y_1; and $\left\{\frac{\beta e}{x_1^2}^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}} + \frac{dEV_2}{dx_1}\right\}$ is the marginal utility of x_1. The marginal utility of x_1 is made up of the direct effect of the good on current utility and the indirect effect of the informational effect on expected utility next period.

Then it can be seen that (5.10) says that the ratio of marginal utilities should equal the ratio of prices for an optimum.

Substitute into budget constraint for y_1:

$$\therefore \quad M = \left[\frac{\alpha p_x p_y}{\frac{\beta e}{x_1^2}^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}} + \frac{dEV_2}{dx_1}}\right]^{1/2} + p_x x_1 \tag{5.11}$$

and $$\frac{dEV_2}{dx_1} = (\alpha\beta\ p_x p_y)^{1/2}\, e^{-\frac{\mu_0}{2} + v + \frac{1}{4r}} \cdot \frac{1}{4M}\, \frac{1}{(rx_1 + \phi)}^2$$

Thus equation (5.11) gives an implicit equation for the optimal value of x_1 which maximises the present value of expected utility.

We can now compare this value of x_1, with the value obtained for the static case, where there is no opportunity for learning.

Proposition 5.1 If the consumer's objective function is given by (5.8), and if x_1^0 is the value of the decision variable which maximises the objective in the non-adaptive case, and x_1^* is the optimal decision rule in the adaptive case then $x_1^* \geqslant x^0$.

<u>Proof</u> In the non-adaptive case the objective is:

$$\max_{x_1 y_1} \quad EU_1^0 = -\frac{\alpha}{y_1} - \frac{\beta e^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}}}{x_1} + \lambda \left[M - p_x x_1 - p_y y_1 \right] \qquad (5.12)$$

F.O.C. $$\frac{dEU_1^0}{dy_1} = \frac{\alpha}{y_1^2} - \lambda p_y = 0 \qquad (5.13)$$

and $$\frac{dEU_1^0}{dx_1} = \frac{\beta e^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}}}{x_1^2} - \lambda p_x = 0 \qquad (5.14)$$

Yielding $$x_1^0 = \frac{M}{p_x + \left(p_y p_x \frac{\alpha}{\beta}\right)^{1/2} e^{\frac{\mu_0}{2} - \frac{1}{4\phi} - \frac{1}{4r}}} \qquad (5.15)$$

The value of x_1^0 in (5.15) can be substituted into (5.9).

$$\frac{dEU_1}{dx_1} = \frac{\beta e^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}}}{(x_1^0)^2} + \frac{1}{4M} \frac{(\alpha\beta\, p_x p_y)^{1/2}}{rx_1^0 + \phi)^2} e^{-\frac{\mu_0}{2} + v + \frac{1}{4r}} - \frac{p_x}{p_y (M - p_x x_1^0)^2}$$

But from (5.13) and (5.14):

$$\frac{\beta e^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}}}{(x_1^0)^2} = \frac{p_x \alpha}{p_y (M - p_x x_1^0)^2}$$

$$\therefore \left.\frac{dEU_1}{dx_1}\right|_{x_1^0} = \frac{M}{p_x + (p_y p_x \frac{\alpha}{\beta})^{1/2} \cdot e^{\frac{\mu_0}{2} - \frac{1}{4\phi} - \frac{1}{4r}}} = \frac{1}{4M} (\alpha \beta p_x p_y)^{1/2} e^{-\frac{\mu_0}{2} + v(x_1^0) + \frac{1}{4r}} > 0$$

Thus at x_1^0, $\frac{dEU_1}{dx_1}$ is increasing in x_1, and we would expect x_1^* to lie further to the right

i.e. $x_1^* > x_1^0$ <u>Q.E.D.</u>

<u>Proposition 5.2</u> If the consumer's objective function is given by (5.8), then there exists a unique pair of decision variables x_1^* and y_1^* which maximise the objective.

3. Comparative Statics

<u>Proposition 5.3</u> A consumer will have a higher level of expected utility, the less risky is the initial prior distribution of the unknown parameter. Risky again being defined as a high initial variance, which is the inverse of the degree of precision.

Expected utility increases as the consumer's initial degree of precision increases.

Now consider the effect of a change in the initial degree of precision on the optimal demand for the new good in the first period:

Proposition 5.4 x_1^* is monotonically non-increasing in ϕ

Thus as the degree of confidence increases, and hence the risk of a false mean falls, the consumer purchases less of the new good.

This is because the higher is the degree of precision, the more confident the consumer is that the prior mean is the true mean, the less the value of information and consequently there is less need to sample and hence x_1^* falls. Consumers demand more information about unfamiliar products.

Consider the effect of a change in the initial advertising statement $\ddot{\mu}_0$, on the demand for the new good in the first period

<u>Proposition 5.5</u> x_1^* is monotonically non-increasing in $\ddot{\mu}_0$.

<u>Proof</u> $$\frac{dx_1^*}{d\mu_0} = -\frac{\dfrac{d^2EU_1}{dx_1 d\mu_0}}{\dfrac{d^2EU_1}{dx_1^2}}$$

We know from the second order conditions : $\dfrac{d^2EU_1}{dx_1^2} < 0$

From (5.10)

$$\frac{d^2EU_1}{dx_1 d\mu_0} = -\frac{\beta e}{x_1^2}^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}} - \frac{1}{8M} \cdot \frac{(\alpha\beta p_x p_y)^{1/2}}{(rx_1 + \phi)^2} \cdot e^{-\frac{\ddot{\mu}_0}{2} + v + \frac{1}{4r}} < 0 \qquad (5.16)$$

<u>Q.E.D.</u>

As $\ddot{\mu}_0$ increases the optimal purchase of the new good falls. It can be seen that the effect of an increase in $\ddot{\mu}_0$ is non-increasing both on the first period consumption of x_1 and the indirect utility function for the second period.

Proposition 5.5 appears counter intuitive, since we might expect that as the expected value of the random variable increases, reflecting an expected increase in product quality, the demand for the good would also increase. However, this proposition shows that this is not the case.

Equation (5.16) is made up of two parts: the direct effect of a change in μ_0 on current utility and an indirect effect through changing the informational aspect in the next period.

Consider firstly the direct effect. We know from equation (11) in Chapter 2 that the effect of a parameter change on the demand for a good can be compared with the income and substitution effects of an implicit price change.

If we write out equation (5.1) as a single period problem in its characteristics form:

$$\max_{c,\,y} \quad U = -\frac{\alpha}{y} - \frac{\beta}{c}$$

$$\text{s.t.} \quad p_y y + \pi_c c = M$$

where $\quad c = \mu x, \quad \pi_c = \dfrac{px}{\mu}$ and $\mu = e^{\mu_0 + \frac{1}{2\phi}}$

This yields the optimal demand for the characteristic

$$c = \frac{M}{\pi_c + \sqrt{p_y \pi_c \frac{\alpha}{\beta}}} \tag{5.17}$$

The effect of a change in the implicit price on the demand for the characteristi can be found by differentiating equation (5.17)

$$\frac{dc}{d\pi c} = -\frac{M}{\left[\pi_c + \sqrt{p_y \pi_c \frac{\alpha}{\beta}}\right]^2} \cdot \left\{1 + \frac{1}{2}\sqrt{\frac{p_y \alpha}{\pi_c \beta}}\right\} \tag{5.18}$$

Substituting (5.18) into (2.11) gives

$$\frac{dx_1}{d\mu} = \frac{1}{\mu}\left\{\frac{p_x}{\mu^2} \cdot \frac{M}{\left[\frac{p_x}{\mu} + \sqrt{\frac{p_y p_x \alpha}{\mu\beta}}\right]^2} \cdot \left[1 + \frac{1}{2}\sqrt{\frac{p_y \mu\alpha}{p_x \beta}}\right]\right\} - \frac{M}{\mu\left[p_x + \sqrt{p_y p_x \frac{\alpha\mu}{\beta}}\right]}$$

$$\therefore \quad \frac{dx_1}{d\mu} \gtrless 0 \quad \text{as } p_x \left[1 + \frac{1}{2}\sqrt{\frac{p_y}{p_x}\frac{\alpha}{\beta}\mu}\right] \gtrless p_x + \sqrt{p_y p_x \frac{\alpha}{\beta}\mu}$$

i.e. $0 < \frac{1}{2}\sqrt{p_y p_x \frac{\alpha}{\beta}\mu}$

$$\therefore \frac{dx_1}{d\mu} < 0$$

Thus we can see that the reason for the inverse relationship between product quality and demand, is due to the second term in equation (2.11). The structure of the utility function is such that an increase in product quality means that the same level of utility can be achieved by purchasing less of the good.

The indirect effect of a change in μ_0 on x_1^*: the second term in equation (5.16) can be explained as follows. Differentiate (5.5) with respect to $\bar{\varepsilon}$, the message provided by the consumer's experience:

$$\frac{dx_2}{d\bar{\varepsilon}} = \frac{Mx_1 r}{rx_1 + \phi}\left\{\frac{-1}{\left(p_x p_y \frac{\alpha}{\beta}\right)^{1/2} e^{\frac{\mu_1}{2} - \frac{1}{4}\frac{1}{\phi + x_1} - \frac{1}{2r}} + 2p_x + \dfrac{p_x^2}{\left(p_x p_y \frac{\alpha}{\beta}\right)^{1/2} e^{\frac{\mu_1}{2} - \frac{1}{4}\frac{1}{\phi + x_1 r} - \frac{1}{2r}}}}\right.$$

But as $\mu_0 \to \infty$; $\frac{dx_2}{d\bar{\varepsilon}} \to 0$. If we define $dx_2/d\bar{\varepsilon}$ as the message effect, then the message effect is reduced as μ_0 increases. Thus the higher the level of μ_0, the more insensitive is the demand for the new good in the second period to the experience actually observed. But if commodity choices become independent of the information obtained from sampling, then the consumer will not bother sampling, and x_1^* is reduced.

Proposition 5.6 x_1^* is monotonically non-increasing in r.

This proposition suggests that an increase in the underlying distribution generating the random variables more widely dispersed. The consumer will have to take more observations in order to hold the same degree of confidence about the subjective mean of the distribution.

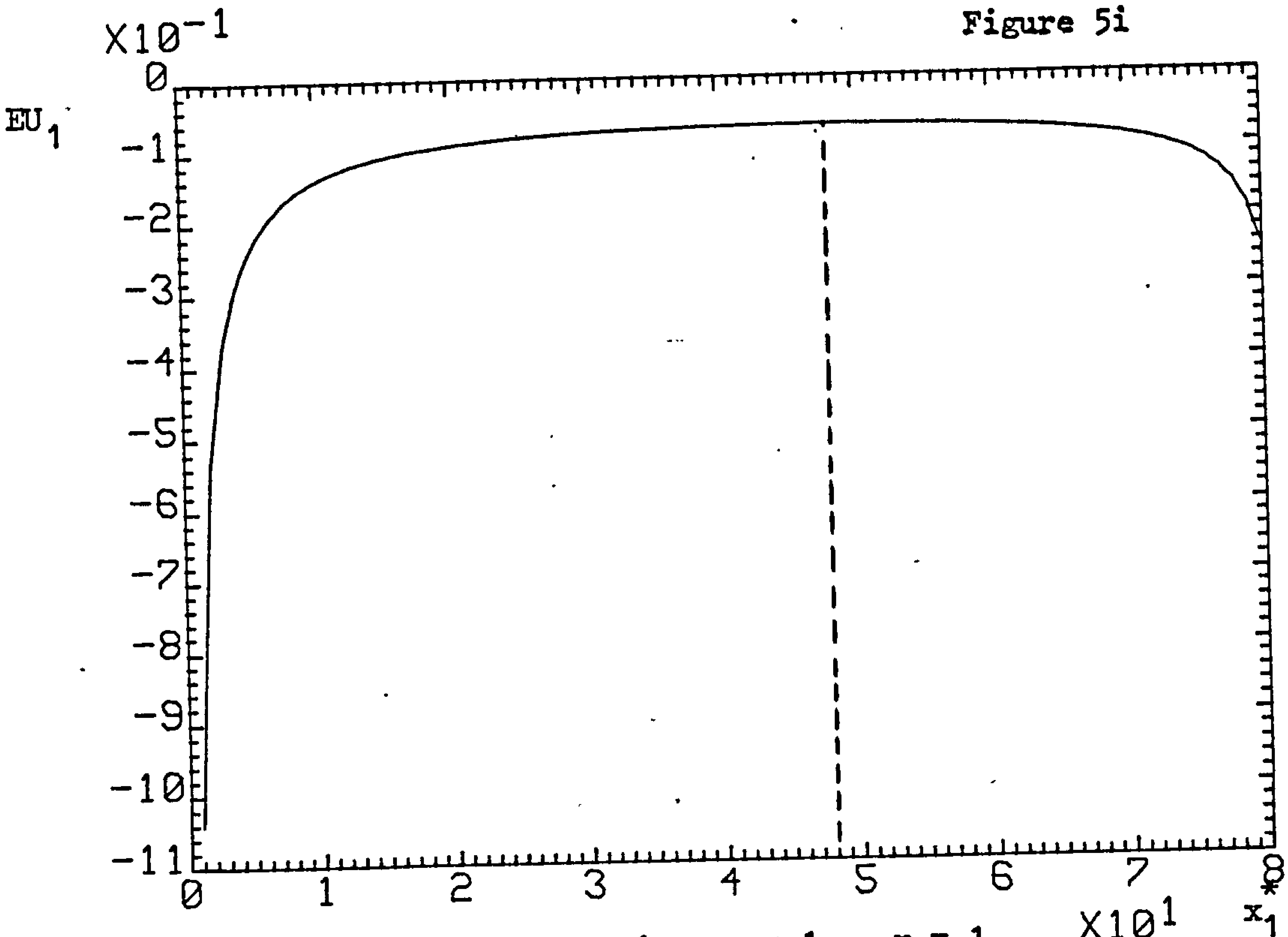
observations; the optimal sample size increases.

The table below and the figures illustrate these comparative static results. There is an internal solution for all the parameter values, which makes sense with convex preferences. Comparing (ii) with (i), we see that an increase in β, increases the demand for x_1. This shows that an increased preference for the new good causes the consumer to purchase more. Row (iii) and figure 5iii illustrate proposition 5.5: an increase in μ_0, reduces the demand for the new good. A reduction in the price of the new good not surprisingly increases the demand for it as shown in (iv). Finally we consider changes in the variance. A reduction in the consumer's degree of confidence about the value of μ_0 being the true value, increases the demand for x_1^*. The reduction in ϕ in (v) also reduces expected utility. This effect of ϕ on x_1^* and EU_1 can be contrasted with propositions 4.4 and 4.5. In the last row and figure (vi) an increase in the underlying degree of precision r, reduces the demand for the new good. Thus an increase in the underlying variance, increases the demand as shown in proposition 5.6.

Table 5.1

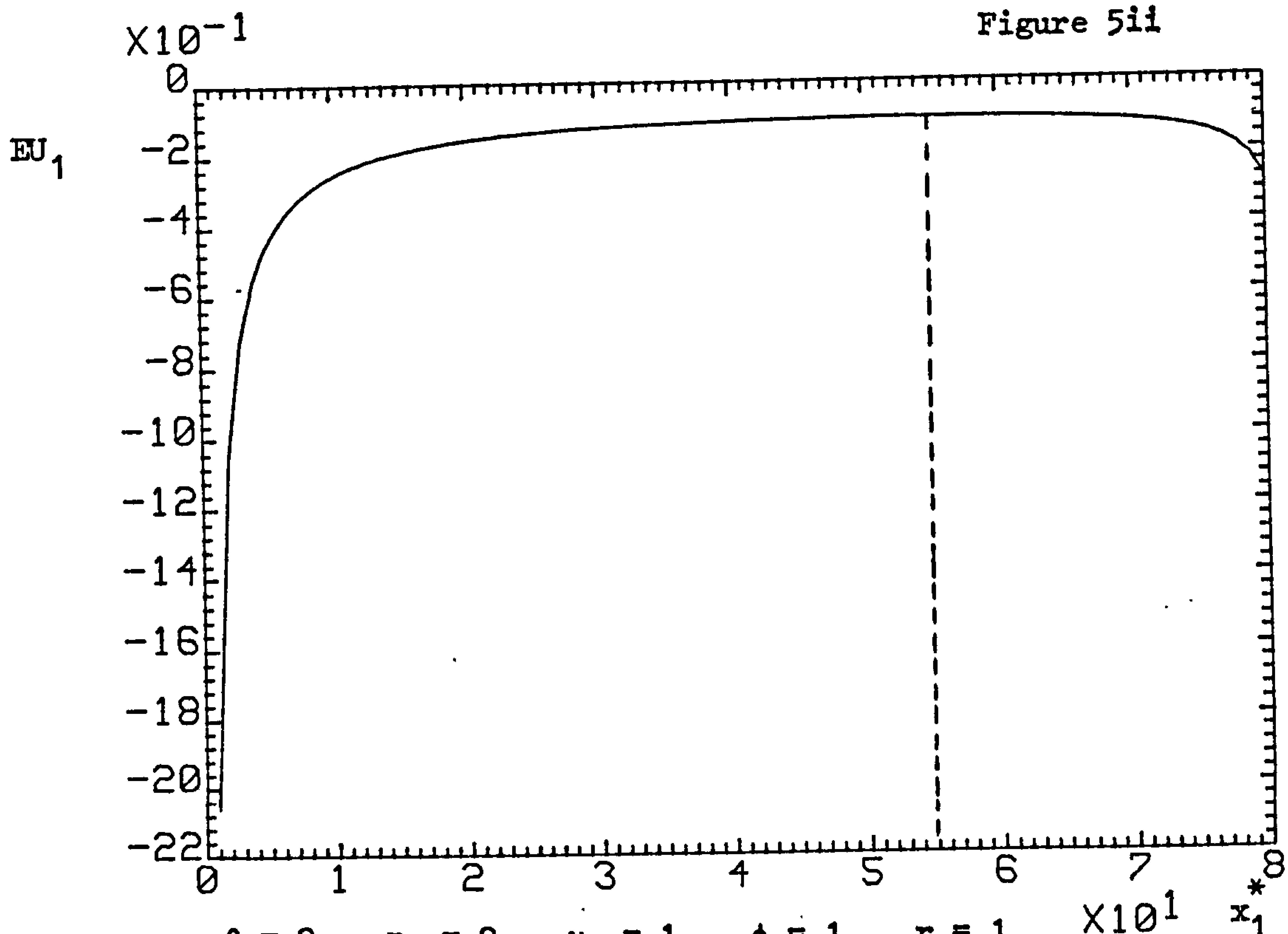
	α	β	p_x	M	μ_0	ϕ	r	x_1^*	EU at x_1^*
(i)	1.0	1.0	2.0	165	1.0	1.0	1.0	48	-0.06867
(ii)	1.0	2.0	2.0	165	1.0	1.0	1.0	55	-0.10629
(iii)	1.0	1.0	2.0	165	5.0	1.0	1.0	13	-0.01695
(iv)	1.0	1.0	1.0	165	5.0	1.0	1.0	20	-0.01544
(v)	1.0	1.0	2.0	165	1.0	1.0	0.1	77	-2.36939
(vi)	1.0	1.0	2.0	165	1.0	10.0	1.0	44	-0.05482

Figure 5i



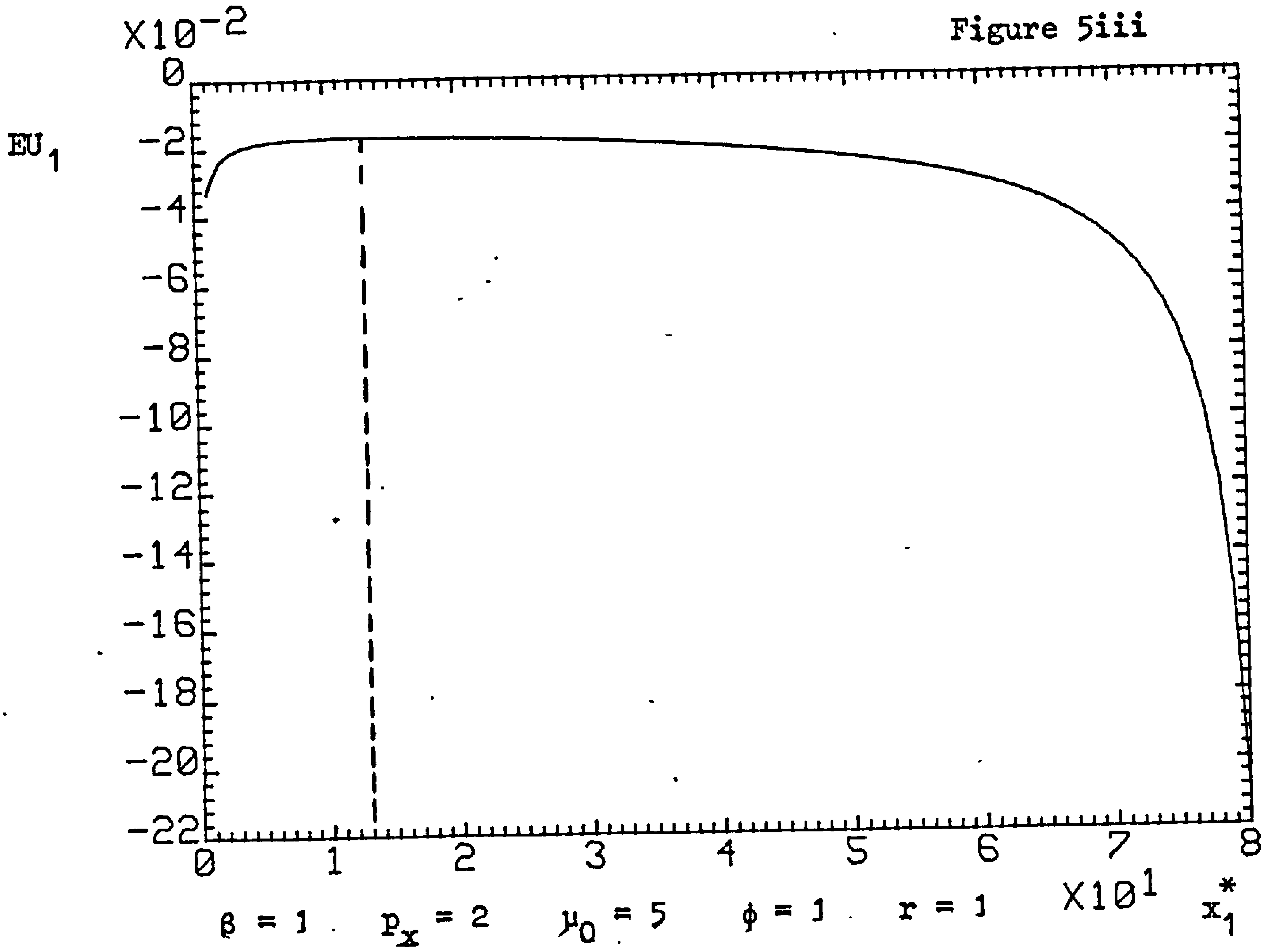
$\beta = 1 \quad p_x = 2 \quad \mu_0 = 1 \quad \phi = 1 \quad r = 1$

Figure 5ii



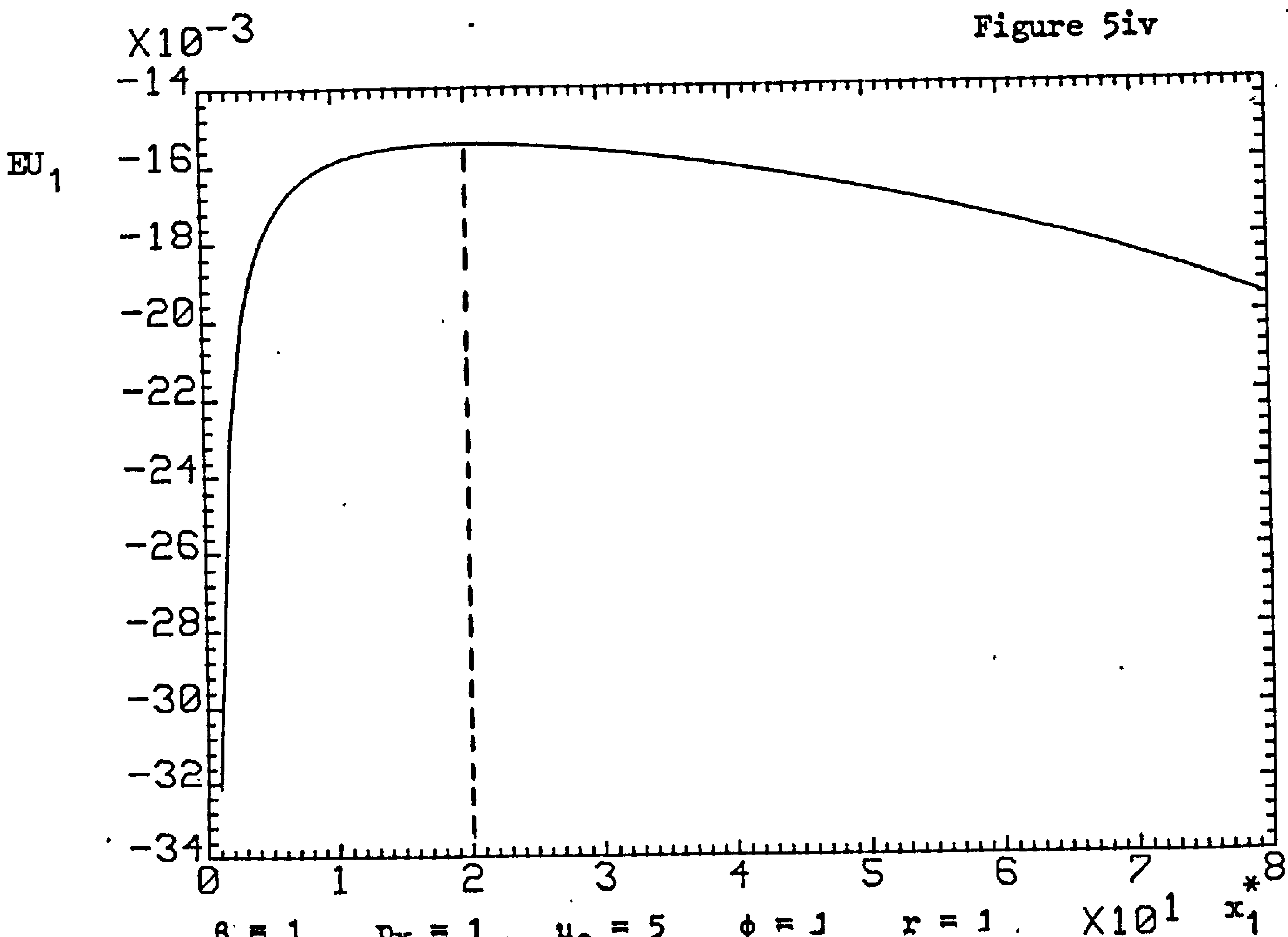
$\beta = 2 \quad p_x = 2 \quad \mu_0 = 1 \quad \phi = 1 \quad r = 1$

Figure 5iii



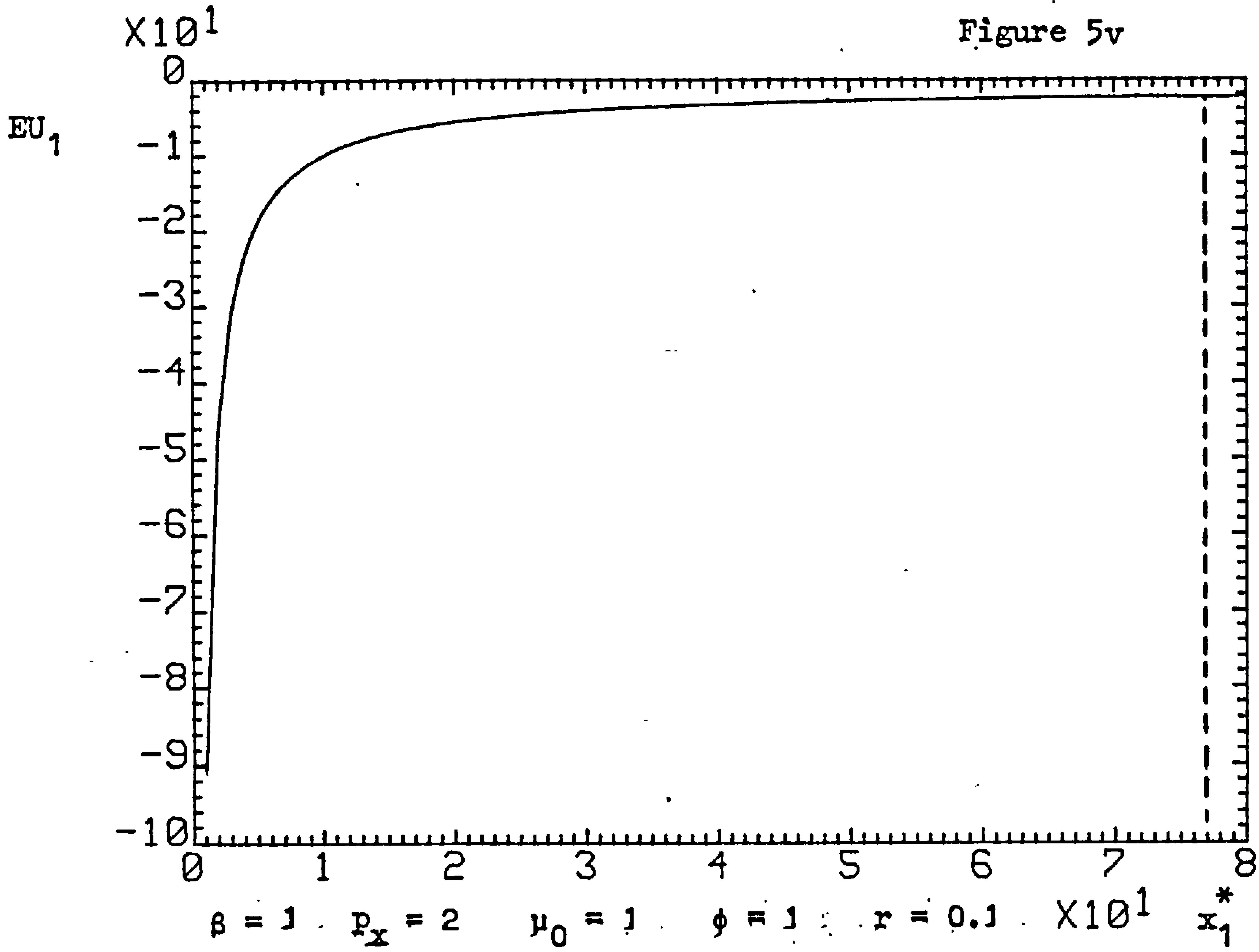
$\beta = 1$ $p_x = 2$ $\mu_0 = 5$ $\phi = 1$ $r = 1$

Figure 5iv



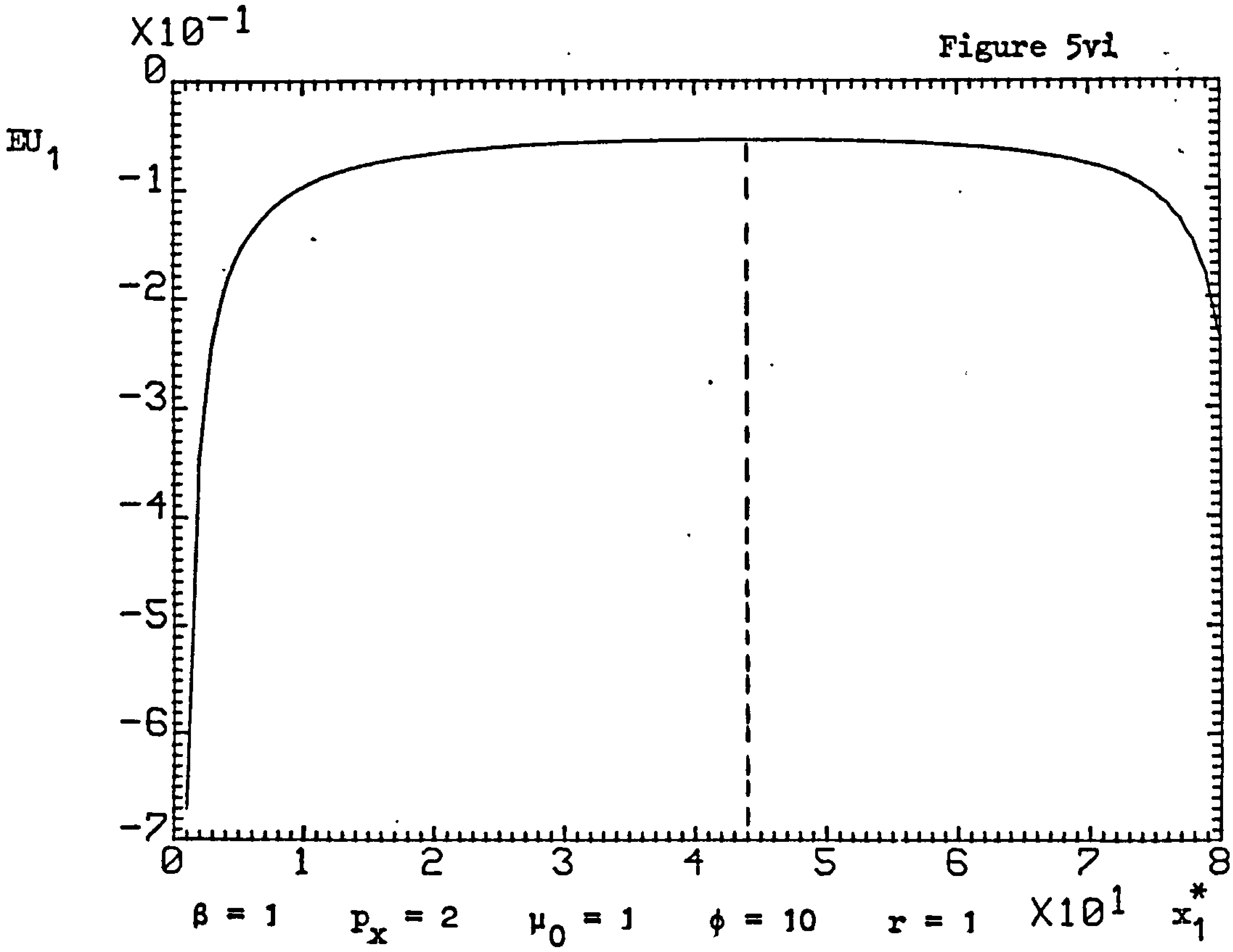
$\beta = 1$ $p_x = 1$ $\mu_0 = 5$ $\phi = 1$ $r = 1$

Figure 5v



$\beta = 1$ $p_x = 2$ $\mu_0 = 1$ $\phi = 1$ $r = 0.1$

Figure 5vi



$\beta = 1$ $p_x = 2$ $\mu_0 = 1$ $\phi = 10$ $r = 1$

4. Conclusions

This chapter has included a concave utility function and a likelihood function which is skewed. The previous chapter considered a linear utility function, and consequently for particular parameter values generated corner solutions such that the consumer purchased one good or the other. Here though, the introduction of convex preferences, means that the consumer always purchases a mixture of the two goods. As previously, the opportunity for learning means that the consumer purchases more of the new good than if the ability to learn did not exist. The importance of the distribution generating the random variables being skewed is evidenced by proposition 5.6: as the degree of skewness increases the consumer purchases more of the new good; he has to sample more. The effect of an increase in the subjective degree of precision on the demand for the good is exactly opposite for the case of the C.E.S. utility function and the linear one.

Appendices

Proposition 5.2, Proof.

$$\frac{d^2EU_1}{dy_1^2} = -\frac{2\alpha}{y_1^3} < 0$$

and $$\frac{d^2EU}{dx_1^2} = -\frac{2\beta e^{-\mu_0+\frac{1}{2\phi}+\frac{1}{2r}}}{x_1^3} + \frac{d^2EV_2}{dx_1^2}$$

and $$\frac{d^2EV_2}{dx_1^2} = -\frac{1}{32M}\frac{(\alpha\beta p_x p_y)^{1/2}}{(rx_1+\phi)^4} \cdot e^{-\frac{\mu_0}{2}+v+\frac{1}{4r}} \cdot \left[1-16(rx_1+\phi)\right] < 0$$

and $$\frac{d^2EU_1}{dy_1^2} \cdot \frac{d^2EU_1}{dx_1^2} > \left(\frac{d^2EU_1}{dx_1 dy_1}\right)^2 \quad \text{since} \quad \frac{d^2EU_1}{dx_1 dy_1} = 0$$

and $$\frac{2\alpha}{y_1^3}\left[\frac{2\beta e^{-\mu_0+\frac{1}{2\phi}+\frac{1}{2r}}}{x_1^3} + \frac{1}{32M} \cdot \frac{(\alpha\beta p_x p_y)^{1/2}}{(rx_1+\phi)^4} e^{-\frac{\mu_0}{2}+v+\frac{1}{4r}} \cdot \left[1-16(rx_1+\phi)\right]\right] > 0$$

Now suppose there are two values of x_1, $\hat{x}_1'$ and $\hat{x}_1''$ that maximise EU_1, then there must be a local minimiser somewhere between $\hat{x}_1'$ and $\hat{x}_1''$. Then the conditions for a minimum are $\frac{dEU_1}{dx_1} = 0$ and $\frac{d^2EU_1}{dx_1^2} > 0$.

But we have shown above that $\frac{d^2EU_1}{dx_1} < 0$, so that minimum can not exist and there is only one maximising value of x_1.

Q.E.D.

Proposition 5.3 Proof.

$$\frac{dEU_1}{d\phi} = \frac{\beta}{x_1} e^{-\mu_0+\frac{1}{2\phi}+\frac{1}{2r}} \cdot \frac{1}{2\phi^2} + \frac{dEV_2}{d\phi}$$

where $$\frac{dEV_2}{d\phi} = -\frac{1}{M}\left\{-\frac{\beta p_x e^{-\mu_0+\frac{1}{2\phi}+\frac{1}{2r}}}{2\phi^2} + 2(\alpha\beta p_x p_y)^{1/2} e^{-\frac{\mu_0}{2}+v+\frac{1}{4r}} \cdot \frac{dv}{d\phi}\right\}$$

where $\frac{dv}{d\phi} = - \frac{\left[\frac{1}{2} x_1^2 r^2 + rx\,\phi + \phi^2\right]}{4\,\phi^2\,(rx_1 + \phi)^2}$

$$\therefore \frac{dEU_1}{d\phi} > 0$$

Q.E.D.

Proposition 5.4, Proof.

$$\frac{dx^*}{d\phi} = - \frac{\frac{d^2EU_1}{dx_1 d\phi}}{\frac{d^2EU_1}{dx_1^2}}$$

Again, we know from the second order conditions that the denominator is negative.

From (5.9)

$$\frac{d^2EU_1}{dx_1 d\phi} = -\frac{1}{2\phi^2} \cdot \beta e^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}} + \frac{d^2EV_2}{dx_1 d\phi}$$

where $\frac{d^2EV_2}{dx_1 d\phi} = \frac{(\alpha\beta p_x p_y)^{1/2}}{4M(rx_1 + \phi)^4}\, e^{-\frac{\mu_0}{2} + \frac{1}{4r} + v} \cdot \left\{(rx_1 + \phi)^2 \frac{dv}{d\phi} - 2(rx_1 + \phi)\right\}$

$$= - \frac{(\alpha\beta p_x p_y)^{1/2}}{4M\,(rx_1 + \phi)^4}\, e^{-\frac{\mu_0}{2} + v + \frac{1}{4}} \cdot \left[\frac{\frac{1}{2} x_1^2 r^2 + r\phi x_1 + \phi^2 + 2(rx_1 + \phi}{4\phi^2}\right.$$

$$\frac{d^2EU_1}{dx_1 d\phi} < 0 \implies \frac{dx^*}{d\phi} < 0$$

Q.E.D.

Propostion 5.6, Proof

Differentiate equation (5.9) w.r.t. r

$$\frac{d^2EU_1}{dx_1dr} = -\frac{\beta e^{-\mu_0 + \frac{1}{2\phi} + \frac{1}{2r}}}{2x_1^2r_2} + \frac{(\alpha\beta p_x p_y)^{1/2}}{4M} \cdot \frac{e^{-\frac{\mu_0}{2} + v + \frac{1}{4r}}}{(rx_1 + \phi)^4}$$

$$\left\{(rx_1 + \phi)^2 \left[\frac{dv}{dr} - \frac{1}{4r^2}\right] - 2(rx_1 + \phi)x_1\right\}$$

and $\frac{dv}{dr} = -\frac{1/2 . x_1}{4(rx_1 + \phi)^2}$

So $\frac{d^2EU_1}{dx_1dr} < 0$, which means $\frac{dx_1^*}{dr} < 0$.

Chapter 6

Learning with a Non-Symmetric Prior

1. Introduction

In this chapter we again consider the implications of the ability to learn for consumer behaviour, but this time we concentrate on changing the structure of beliefs. We revert back to using a linear utility function, but suppose that the consumer's beliefs about the value of the unknown parameter are not symmetrically distributed about the value given by the advertising message. Instead it is assumed that the consumer believes that the advertising message states the most optimistic view of product quality, and that the actual level of product quality is less than this stated value. In Chapter 5, we assumed the likelihood function was not symmetric, now we suppose the prior distribution is non-symmetric. Consumer's distrust of advertising accuracy is built into the subjective distribution. We are interested in how behaviour is affected as this distrust becomes more acute.

2. Derivation of the Posterior

Consider a sceptical consumer with a linear utility function

$$E\, u_t = y_t + E(\varepsilon)\, x_t$$

The distribution of ε is known by the consumer to be uniform with parameters (0, z), but the value of z is not known.

$$E(\varepsilon) = \int_{\forall\varepsilon} \varepsilon\, f(\varepsilon)\, d\varepsilon$$

where $$f(\varepsilon) = \begin{cases} \frac{1}{z} & 0 \leqslant \varepsilon \leqslant z \\ 0 & \text{otherwise} \end{cases}$$

$$\therefore \quad E\, u_t = y_t + \frac{z}{2}\, x_t$$

The consumer's subjective beliefs about the prior distribution of z are given by f(z) which is a form of the Inverse Pareto Distribution with parameters (a, b).

$$f(z|a, b) = \begin{cases} \frac{(a+1)\, z^a}{b^{a+1}} & \text{for } 0 < z < b \\ 0 & \text{otherwise} \end{cases}$$

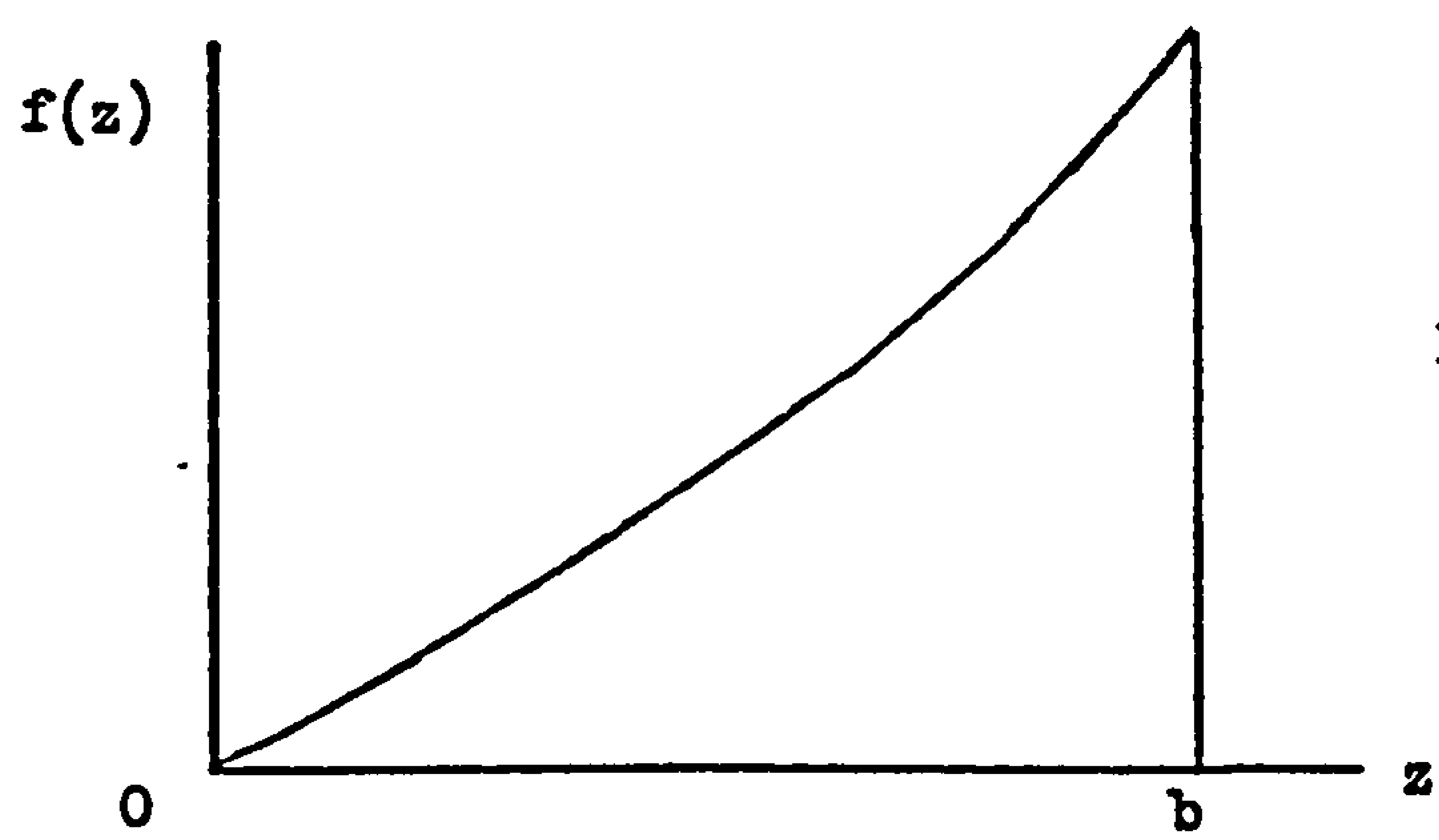


Diagram 6.1

<u>Lemma 6.1</u>: To show that f(z) is a proper density function

$$\int_{\forall z} f(z)\, dz = 1$$

<u>Proof</u>: $$\int f(z)\, dz = \int_0^b \frac{(a+1)\, z^a}{b^{a+1}}\, dz$$

$$= \frac{a+1}{b^{a+1}} \cdot \left[\frac{z^{a+1}}{a+1}\right]_0^b$$

$$= \frac{a+1}{b^{a+1}} \cdot \frac{b^{a+1}}{a+1}$$

$$= 1$$

<u>Q.E.D.</u>

In fact the Inverse Pareto Distribution is a simple proportional transformation of the beta distribution with parameters $(a + 1, 1)$.[1]/

We suppose that b is a parameter in the utility function supplied by the firm in the form of an advertising statement. Consequently the consumer is sceptical about the accuracy of b being the true value of the unknown parameter in the uniform distribution. The higher is b the higher is the mean of the uniform distribution.

But although the consumer is dubious about the accuracy of the advertising statement, he still believes that the true value of z is somewhere in the vicinity of b, though to the left. Thus we arrive at the skewed distribution. The other parameter a can again be thought of as a measure of the consumer's degree of confidence in the value of b representing the true value of z; since low values of a mean that the distribution is more spread out than a large value of a.

As previously in a two period problem we need to derive the posterior distribution of the unknown parameter z.

Consider first the mean of the prior distribution.

<u>Lemma 6.2</u>: The expected value of a random variable with an Inverse Pareto Distribution and parameters (a, b) is $b \left(\frac{a+1}{a+2}\right)$

Proof: $E(z) = \int_0^b \frac{a+1}{b^{a+1}} \cdot z^{a+1}\, dz$

i.e. $= b\left(\frac{a+1}{a+2}\right)$

Q.E.D.

Lemma 6.3: The posterior distribution of z is given by

$$g(z|\cdot) = \frac{(a-x+1)}{b^{a-x+1} - b_0^{a-x+1}} \cdot z^{a-x} \qquad b_0 < z \leq b$$

where $b_0 = \max(\varepsilon_1, \ldots\ldots \varepsilon_x)$

Proof:

With each unit of x that is consumed, the consumer observes the quantity of the characteristic : the consumer observes $(\varepsilon_1 \ldots\ldots \varepsilon_x)$. The distribution of each observation given the parameter z is:

$$f(\varepsilon_i|z) = \begin{cases} \frac{1}{z} & \text{for } \varepsilon_i \leq z \\ 0 & \text{otherwise} \end{cases}$$

since the consumer knows that each ε_i comes from a uniform distribution (0, z). If each of the ε_i are independently and identically distributed then the joint distribution can be written as:

$$f(\varepsilon_1 \ldots\ldots \varepsilon_x|z) = f(\varepsilon_1|z)\, f(\varepsilon_2|z) \ldots\ldots f(\varepsilon_x|z)$$

$$= \begin{cases} \frac{1}{z^x} & \forall\, \varepsilon_i \leq z \\ 0 & \text{otherwise} \end{cases}$$

But if each $\varepsilon_i < z$ then $\max \varepsilon_i \leq z$, $f(\varepsilon_i \ldots\ldots \varepsilon_x|z)$ is the likelihood function of the sample, given the parameter z. Bayes' Theorem states:

$$g(z|.) \propto f(\varepsilon_1 \ldots\ldots \varepsilon_x|z)\, f(z|a, b)$$

$$\propto \frac{1}{z^x} \cdot \frac{(a+1)}{b^{a+1}} \cdot z^a \quad \text{for } 0 < z < b \text{ and } b_0 < b$$

$$\therefore \; g(z|.) \propto \frac{a+1}{b^{a+1}} \cdot z^{a-x} \qquad b_0 < z < b$$

Making the posterior distribution into a proper distribution function

$$g(z|\cdot) = \frac{(a - x + 1)}{b^{a - x + 1} - b_0^{a - x + 1}} \cdot z^{a - x} \qquad b_0 < z < b$$

Q.E.D.

Lemma 6.4: To show that $g(z|\cdot)$ is a proper density function

$$\int_{b_0}^{b} g(z|\cdot)\, dz = 1$$

Proof:
$$\int_{b_0}^{b} \frac{(a - x + 1)}{b^{a - x + 1} - b_0^{a - x + 1}} \cdot z^{a - x}\, dz$$

$$= \frac{a - x + 1}{b^{a - x + 1} - b_0^{a - x + 1}} \left[\frac{z^{a - x + 1}}{a - x + 1}\right]_{b_0}^{b}$$

Evaluating the definite integral

$$= \frac{(a - x + 1)}{b^{a - x + 1} - b_0^{a - x + 1}} \left[\frac{b^{a - x + 1} - b_0^{a - x + 1}}{(a - x + 1)}\right] = 1$$

Q.E.D.

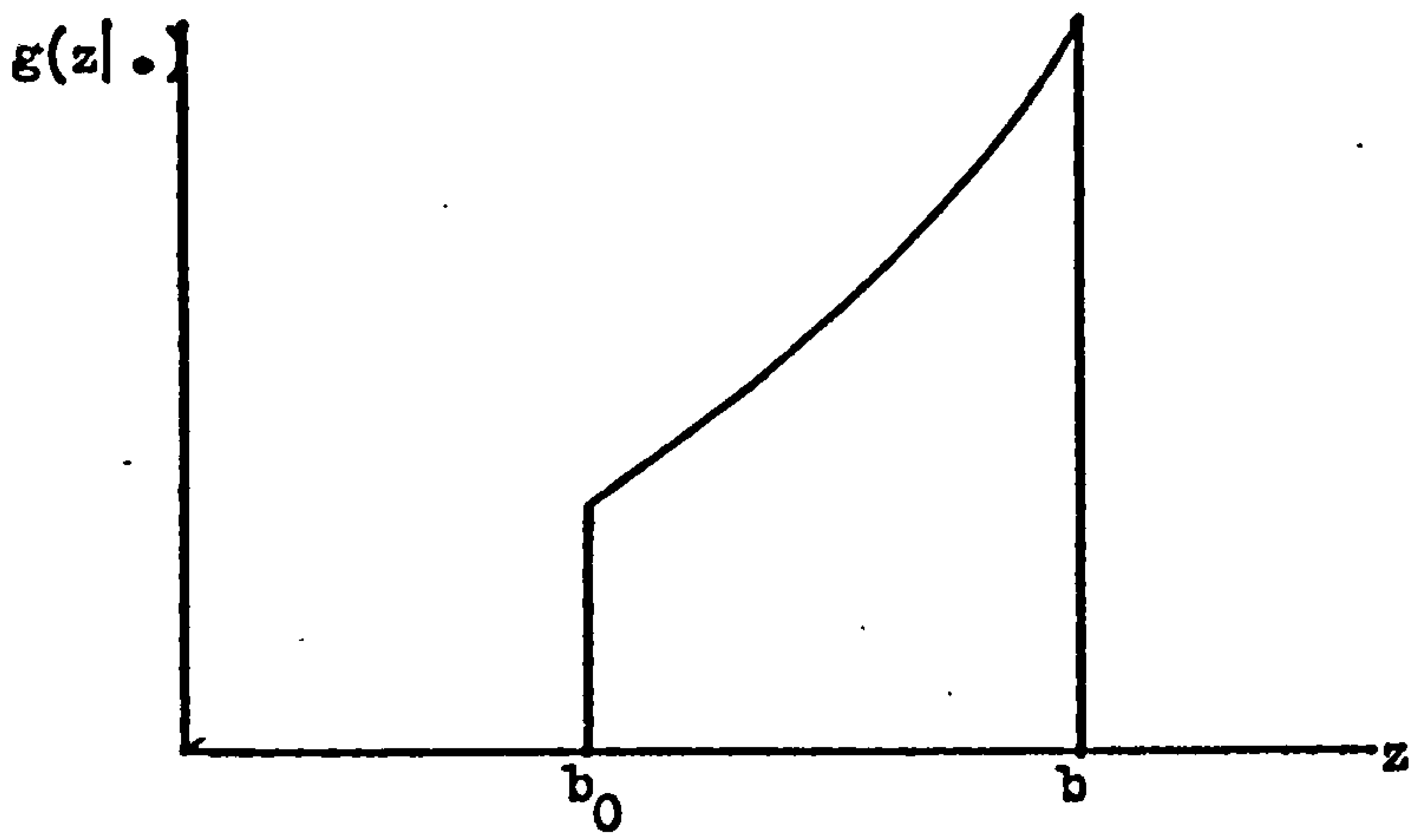


Diagram 6.2

Diagram 6.2 shows the shape of the derived posterior distribution.

Lemma 6.5: The posterior mean of z is given by E(z)

$$E(z) = \frac{a - x + 1}{b^{a - x + 1} - b_0^{a - x + 1}} \cdot \frac{b^{a - x + 2} - b_0^{a - x + 2}}{a - x + 2}$$

Proof: $$E(z) = \int_{b_0}^{b} z\, g(z)\, dz$$

$$= \frac{a - x + 1}{b^{a-x+1} - b_0^{a-x+1}} \int_{b_0}^{b} z^{a-x+1}\, dz$$

$$= \frac{a - x + 1}{a - x + 2} \cdot \frac{b^{a-x+2} - b_0^{a-x+2}}{b^{a-x+1} - b_0^{a-x+1}}$$

Q.E.D.

3. The Model with a Linear Utility Function

Thus the objective faced by a consumer in a two period environment is

$$\max_{y_1, x_1} EU_1 = \sum_{t=1}^{2} \frac{1}{(1+i)^{t-1}} \left[y_t + E_t(z) \frac{x_t}{2} \right] \qquad (6.1)$$

where $E_t(z)$ is the expected value of the unknown parameter, and the time subscript indicates that the expectations operator varies over time. In the first period $E_1(z)$ is given by Lemma 6.2. In the second period the consumer uses the posterior mean derived in Lemma 6.5. But at the start of the first period the value of b_0, the maximum value of the observed levels of product quality consumed in the first period, is still unobserved, and is itself a random variable.

Lemma 6.6: The conditional distribution of b_0 given the value z is

$$\frac{x\, b_0^{x-1}}{z^x}$$

Proof: The distribution of b_0 is required

$$F_{b_0}(\beta) = Pr\left[b_0 < \beta \right] = Pr\left[\varepsilon_1 < \beta,\ \varepsilon_2 < \beta \ldots\ldots \varepsilon_x < \beta \right]$$

since if $b_0 = \max \{ \varepsilon_1 \ldots\ldots \varepsilon_x \}$ then the largest ε_i is less than β only if all the ε_i's are less than β.

$$\therefore \quad F_{b_0}(\beta) = \prod_{i=1}^{x} \Pr\left[\varepsilon_i < \beta\right] = \prod_{i=1}^{x} F_{\varepsilon_i}(\beta) = \left[F_{\varepsilon_i}(\beta)\right]^{x}$$

if it is assumed that the $(\varepsilon_1 \ldots\ldots \varepsilon_x)$ are independently and identically distributed.

$$f_{b_0}(\beta) = \frac{d\, F_{b_0}(\beta)}{d\beta} = x\left[F_{\varepsilon_i}(\beta)\right]^{x-1} f_{\varepsilon_i}(\beta)$$

For a uniform distribution

$$f_{b_0}(b_0) = x\left(\frac{b_0}{z}\right)^{x-1} \cdot \frac{1}{z}$$

$$= \frac{x\, b_0^{x-1}}{z^x} \qquad 0 < b_0 < z$$

<u>Q.E.D.</u>

But Lemma 6.6. only gives the conditional distribution of b_0, which depends on the unknown parameter z. We need to obtain the unconditional distribution.

<u>Lemma 6.7</u>: The unconditional distribution of b_0 is given by

$$\frac{x}{b^x}\left(\frac{a+1}{a-x+1}\right) b_0^{x-1} \qquad 0 < b_0 < b \sqrt[x]{\frac{a-x+1}{a+1}}$$

<u>Proof</u>: Theorem of Total Probabilities states

$$f(b_0) = \int_0^b f(b_0 | z)\, f(z)\, dz$$

$$\therefore \quad = \int_0^b \frac{x\, b_0^{x-1}}{z^x} \cdot \frac{(a+1)\, z^a}{b^{a+1}}\, dz$$

$$f(b_0) = x\left(\frac{a+1}{a-x+1}\right) \cdot \frac{b_0^{x-1}}{b^x}$$

and the range of b_0 necessary to make it into a proper distribution is

$$\left(0,\ b\left(\frac{a-x+1}{a+1}\right)^{\frac{1}{x}}\right).$$

<u>Q.E.D.</u>

The consumer maximises expected utility in the final period with respect to the purchases of the two goods x_2 and y_2; but if we assume a linear budget constraint, then the problem can be written as a function of one variable:

$$\max_{x_2} \; EU_2 = M_2 - px_2 + \frac{E_2(z)}{2} x_2 \qquad (6.2)$$

where $E_2(z)$ is given by Lemma 6.5.

First order conditions yield:

$$\frac{d(EU_2)}{dx_2} = -p + \frac{E_2(z)}{2}$$

$$\therefore \quad \text{If } 2p < E_2(z) \Rightarrow x_2 = \frac{M_2}{p}$$

$$\text{If } 2p > E_2(z) \Rightarrow x_2 = 0$$

Then the expected indirect utility function can be written out as:

$$EV_2 = \int_0^{b_0^*} M_2 \, f(b_0) \, db_0 + \int_{b_0^*}^{b\left(\frac{a-x+1}{a+1}\right)^{1/x}} \frac{M_2}{p} \cdot \frac{E_2(z)}{2} f(b_0) \, db_0 \qquad (6.3)$$

where $\quad b_0^*$ satisfied $2p = \frac{(a - x + 1)}{(a - x + 2)} \cdot \frac{(b^{a-x+2} - b_0^{a-x+2})}{(b^{a-x+1} - b^{a-x+1})}$

and $f(b_0)$ is given by Lemma 6.7.

At the start of the first period the consumer faces the following problem:

$$\max_{x_1} \; EU_1 = M_1 - px_1 + \frac{E_1(z)}{2} x_1 + \frac{1}{1+i} EV_2 \qquad (6.4)$$

where $E_1(z)$ is given by Lemma 6.2, and EV_2 is given by equation (6.3).

First order conditions yield

$$\frac{d\,EU_1}{dx_1} = -p + \frac{b}{2} \cdot \frac{(a+1)}{(a+2)} + \frac{1}{1+i} \frac{dEV_2}{dx_1} \qquad (6.5)$$

The consumer sets (6.5) equal to zero to obtain the optimal x_1^*. As previously the expression will be complex, and the solution can only be found numerically. In fact the expression this time is so complex that the comparative statics' results are only given for numerical values of the parameters. Table 6.1 gives the solutions to some comparative statics' exercises.

Table 6.1 3/

	a	b	p	M	x_1^*	EU_1 at x_1^*
(i)	1.5	1.0	0.36	0.72	1.05	1.44749
(ii)	1.4	1.0	0.36	0.72	0.75	1.44244
(iii)	1.0	1.0	0.34	0.68	0.7	1.36675
(iv)	0.9	1.0	0.34	0.68	0.45	1.36191
(v)	2.0	1.0	0.38	0.76	0.85	1.52095
(vi)	2.0	1.05	0.38	0.76	1.80	1.563125
(vii)	2.0	1.0	0.3805	0.76	0.70	1.52038

Row (i), illustrated by figure 6(i), shows that an internal solution exists. In a single period problem, from Lemma 6.2, for the given parameters, the expected value of the unknown variable is 0.359; so with a linear utility function and $p = 0.36$, the consumer would not purchase any of the new good. However, the positive value of information in this adaptive case means that one unit of x_1 contributes more to utility than just its direct effect, and the consumer does purchase the new good. Row (ii), and figure 6(ii), show the effect of a reduction in the parameter a. This controls the beliefs of the consumer in that the value b provided by the firm is the true upper limit of the underlying uniform distribution. If $a = 0$, then $f(z)$ is a uniform distribution, and the consumer has no confidence in b being the true value. For $0 < a < 1$, $f(z)$ is concave ; $a = 1$, $f(z)$ is linear; and for $a > 1$, $f(z)$ is convex. This is illustrated in diagram 6.3.

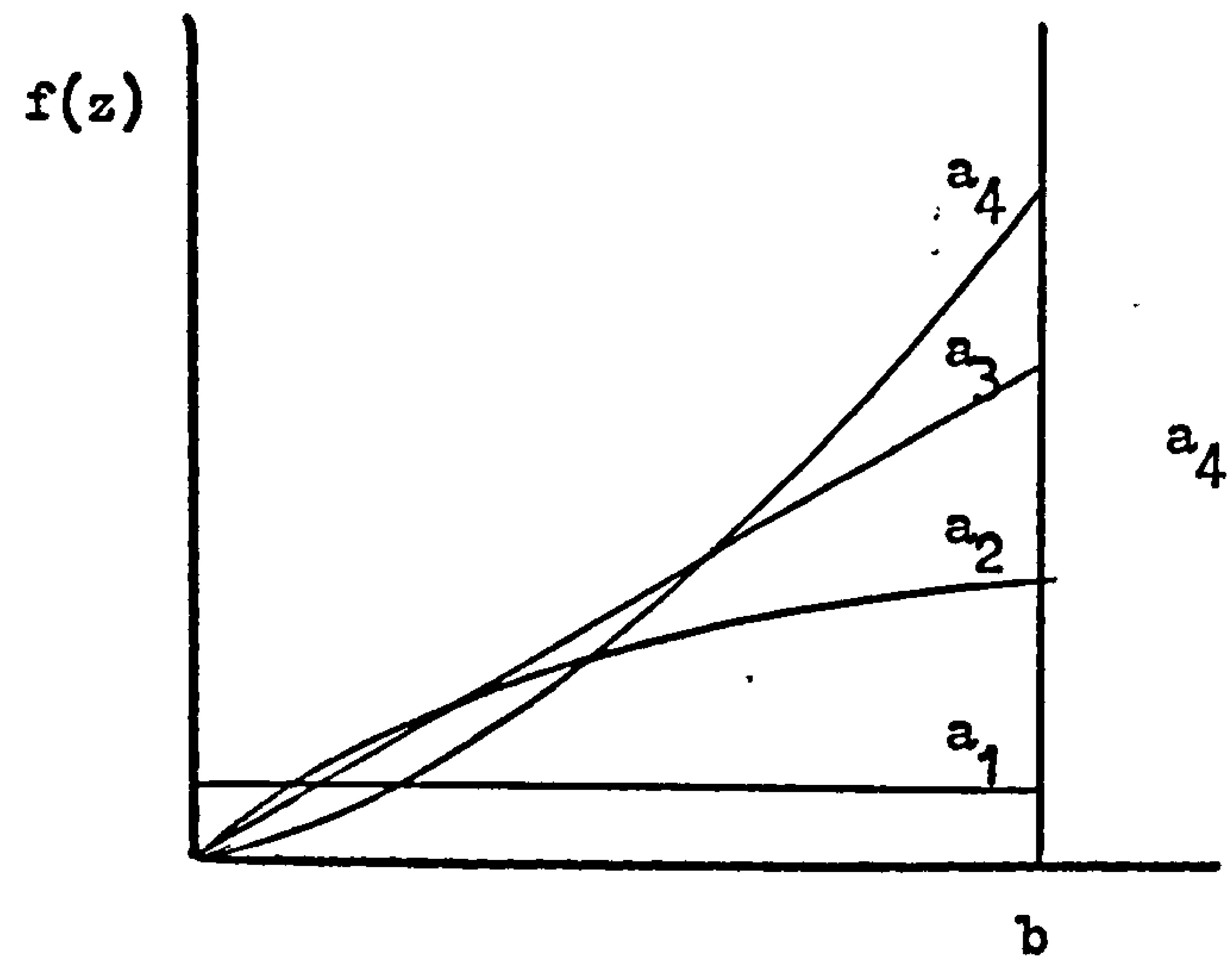


Diagram 6.3

$a_4 > a_3 > a_2 > a_1 = 0$

As a increases, the consumer places greater reliance on the value of b being the true value. Rows (ii), (iii) and (iv) show that an increase in a increases x_1^*. Thus as the consumer becomes more confident of the truthfulness of the advertising message, he purchases more of the new good, and further his expected utility increases.

Row (vi), illustrated by figures 6(iii), and row (vii), illustrated by figure 6(iv), in comparison with the parameter values in Row (v), show that as expected : an increase in b increases x_1^* and EU_1; and an increase in p reduces x_1^* and EU_1.

Figure 6(i)

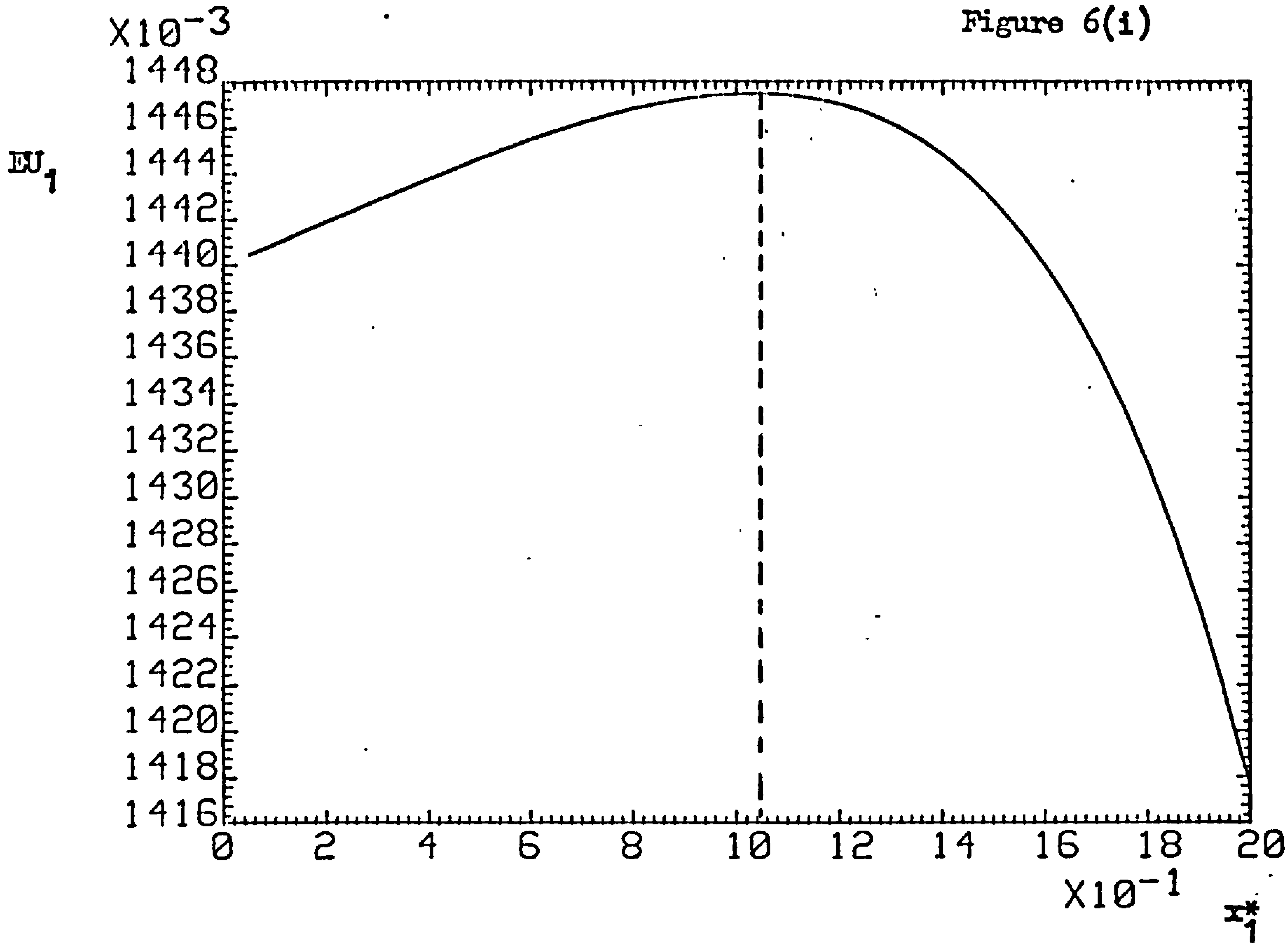


Figure 6(ii)

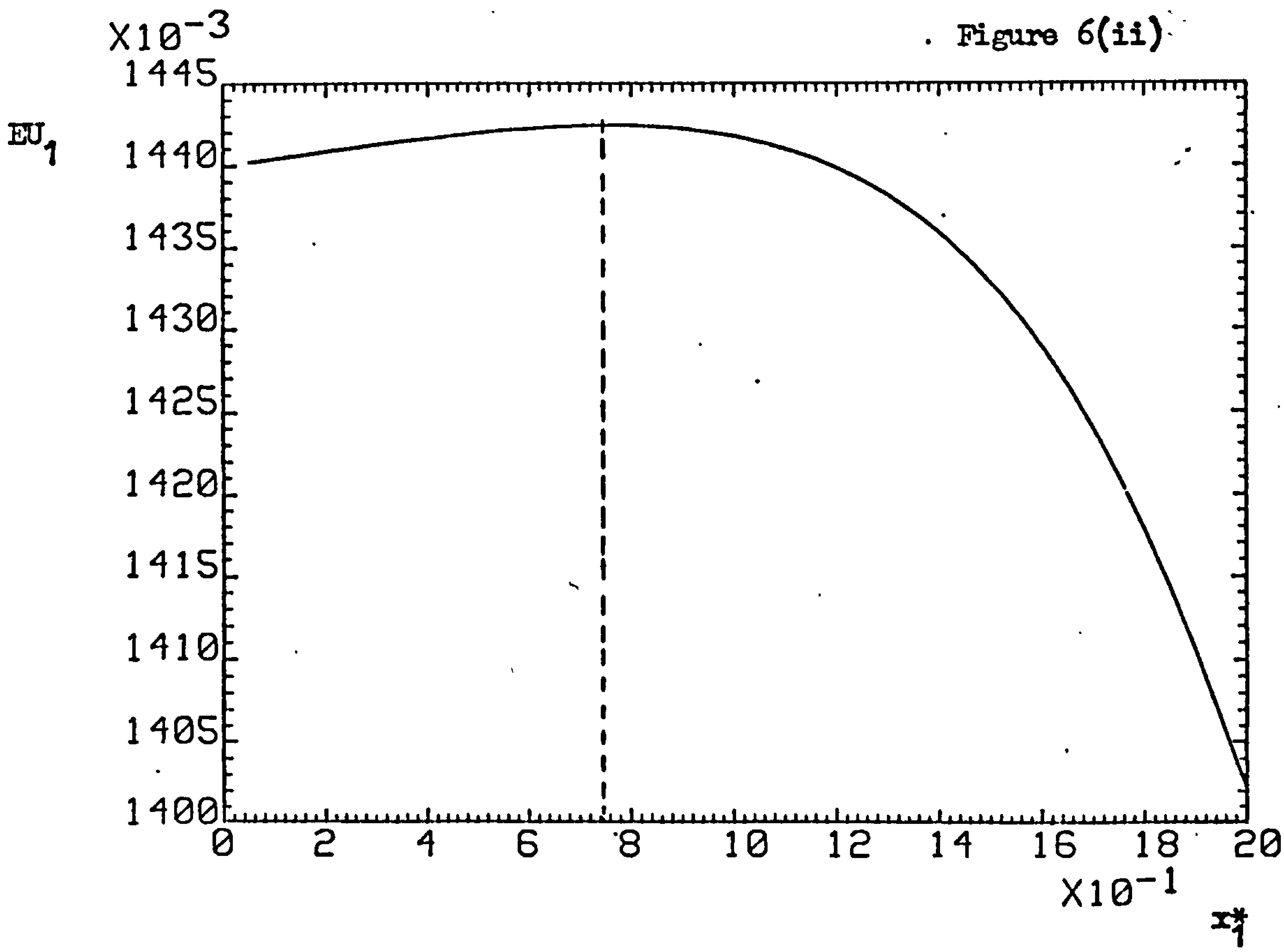


Figure 6(iii)

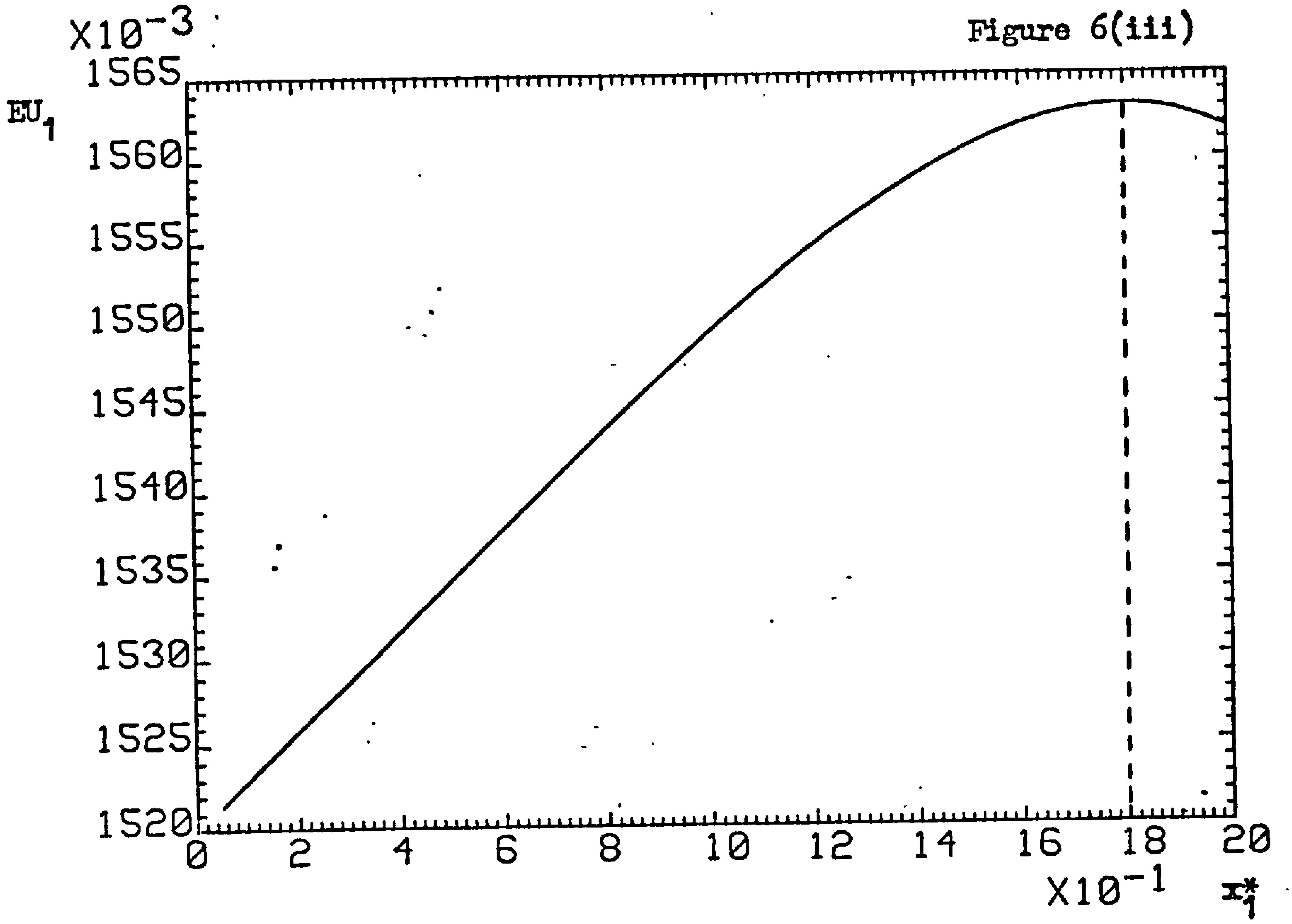
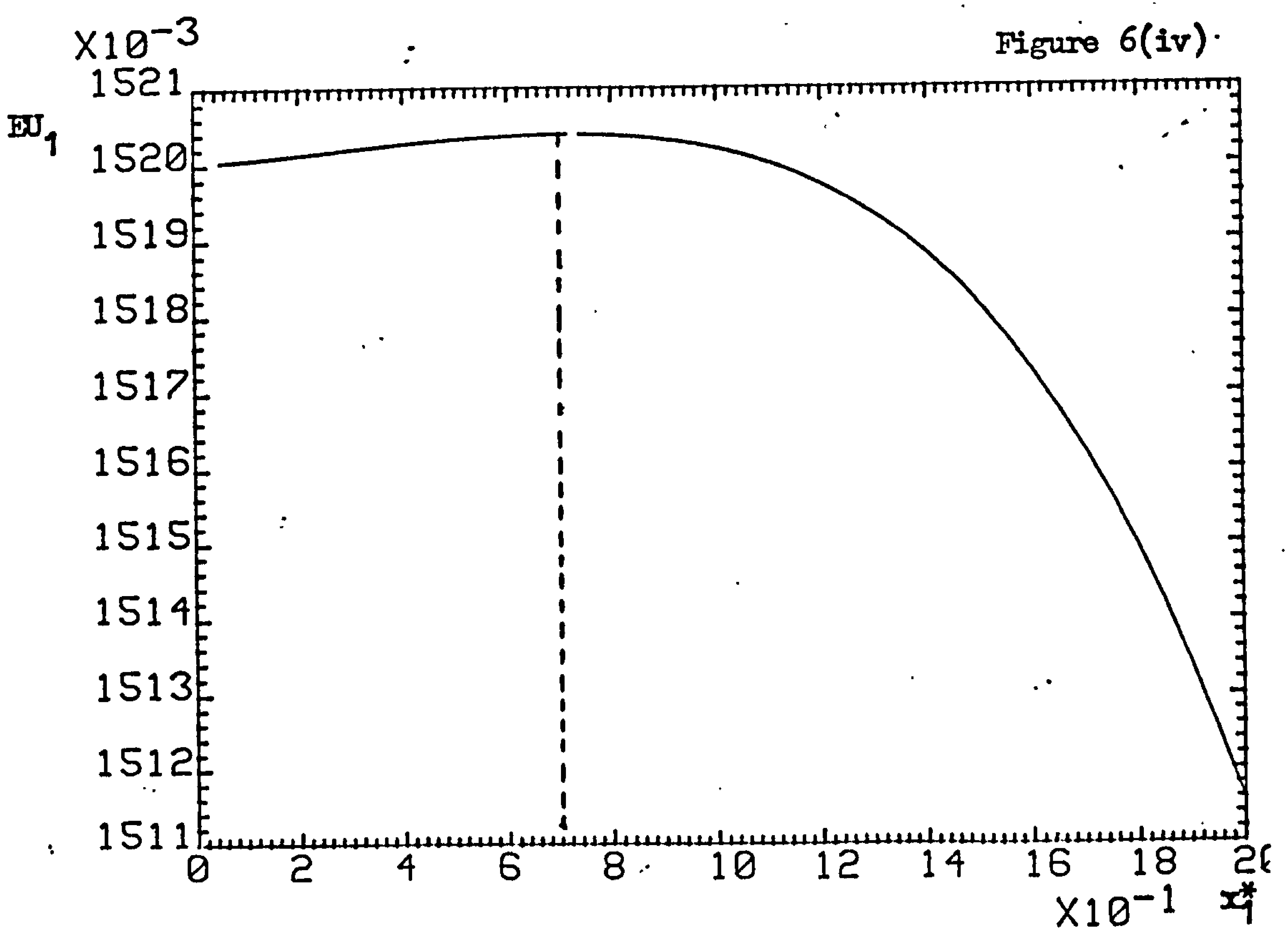


Figure 6(iv)



Conclusion

The results show that in line with Chapter 4, information aquisition means that the decision variable in the first period will have an internal solution, for the distributions considered here. In comparison to Chapter 4, an increase in the parameter a, which represents the consumer's beliefs about the accuracy of the advertising message, increases the value of the decision variable and increases the value of expected utility.

Notes

1/ This is shown in Appendix 6.1

2/ See Appendix 6.2 for the solution to b_0^*

3/ The numerical results are calculated for values of the decision variable between 0 and 2 in intervals of 0.05. This is because if a = 0 then the upper limit in equation (6.3) will be negative if the optimal value of x_1 exceeds unity. Thus x is constrained to take on values in the range (0.2), and consequently the income variable is scaled down accordingly.

Appendix 6.1

To show $f(z) = \dfrac{(1 + a)\, z^{a}}{b^{a + 1}}$ is a transformation of the beta distribution.

The general form of the beta distribution is

$$f(x) = \frac{1}{B(p,\ q)} \cdot x^{p - 1}\,(1 - x)^{q - 1} \qquad 0 \leq x \leq 1$$

where
$$B(p,\ q) = \frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p + q)}$$

Let $p - 1 = a,\ q = 1$

$$\therefore\quad f(x) = \frac{1}{B(1 + a,\ 1)} \cdot x^{a} \qquad 0 \leq x \leq 1$$

and
$$B(1 + a,\ 1) = \frac{\Gamma(1 + a)\,\Gamma(1)}{\Gamma(z + a)} = \frac{\Gamma(1 + a)}{(1 + a)\,\Gamma(1 + a)} = \frac{1}{1 + a}$$

$$\text{i.e. } f(x) = (1 + a)\,x^{a} \qquad 0 \leq x \leq 1$$

Now if
$$x = \frac{z}{b} \qquad \text{so} \qquad 0 \leq z \leq b$$

$$dx = \frac{1}{b}\,dz$$

Then
$$(1 + a)\,x^{a}\,dx = \frac{(1 + a)\,z^{a}}{b^{a + 1}}\,dz$$

which is the inverse Pareto distribution.

Appendix 6.2

In order to solve the polynomial in (6.3) for b_0^* we may make an approximation.

We have:

$$\frac{b^{a-x+2} - b_0^{a-x+2}}{b^{a-x+1} - b_0^{a-x+1}} = 2p\,\frac{(a-x+2)}{(a-x+1)} \qquad (6.6)$$

We approximate b_0^{a-x+i} by expanding b_0^{a-x+i} around the point b.

$$b_0^{a-x+i} \simeq b^{a-x+i} + (a-x+i)(b_0 - b)\, b^{a-x+i-1}$$
$$+ \frac{(a-x+i)}{2}(a-x+i-1)\,(b_0 - b)^2\, b^{a-x+i-2}$$

Making the substituion for b_0^{a-x+2} and b_0^{a-x+1} in (6.3).

$$\therefore\; b_0^* \simeq \frac{2b\,(2p - b)}{b(a-x+1) - 2p\,(a-x)} + b$$

Chapter 7

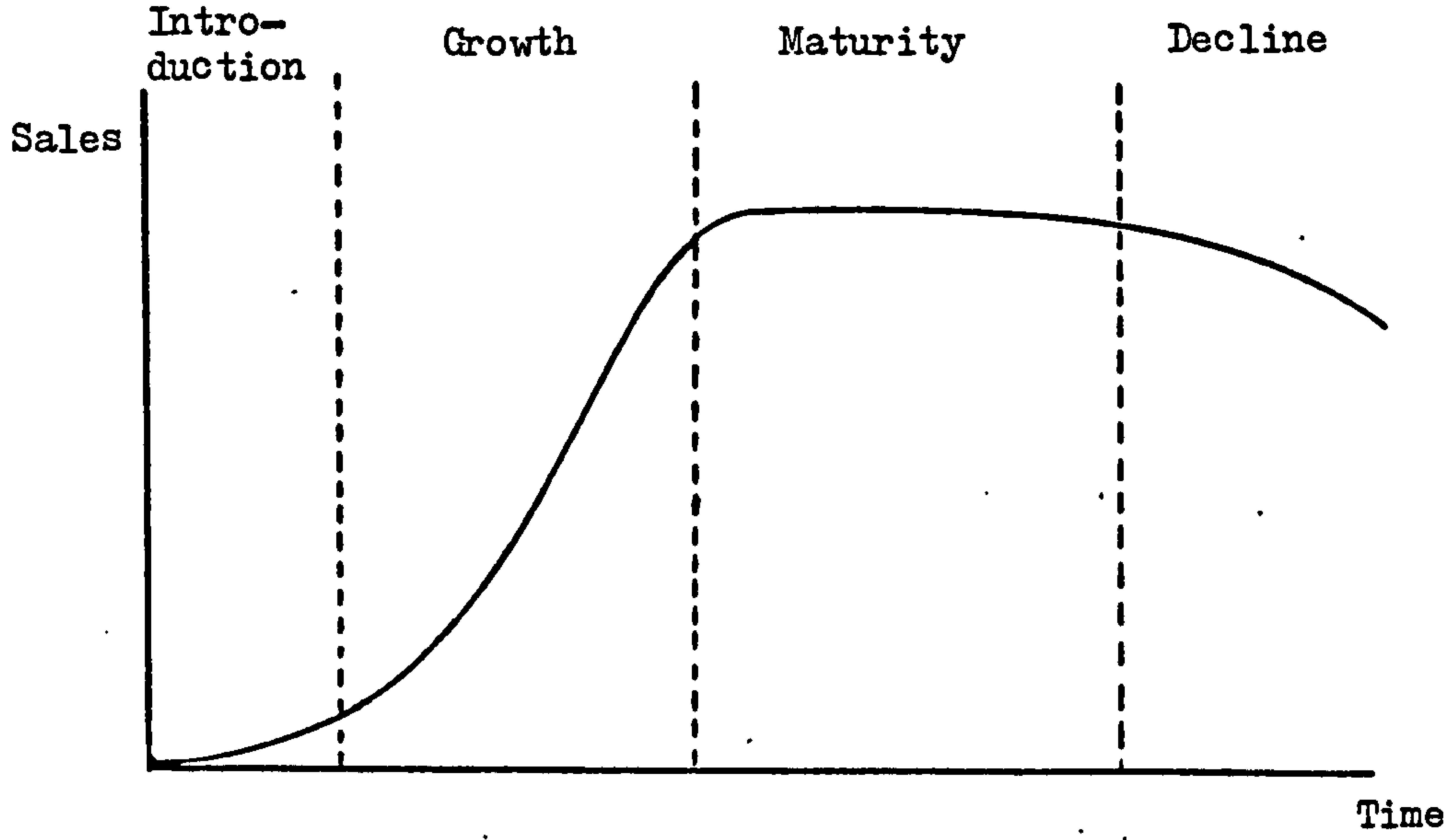
Multi-period Analysis

1. Introduction

In this chapter we wish to extend the results and implications Chapter 4 into a many period environment. Again we consider the introduction of a new product whose quality is unknown by the consumer.

An extension of Chapter 4 would be to construct a forward-looking optimal plan over many time periods. Given his current beliefs, the consumer works out the optimal purchases of the good over all future periods, knowing that his knowledge will accumulate as he progresses. However his particular plan only exists for one period, since the consumer having purchased the optimal first period quantity of the good, changes his beliefs, and constructs a new optimal plan. Of more interest is the value of the consumer's actual purchases, in the first period of each revised optimal plan.

Both the marketing and technological diffusion literatures(1) are concerned with the concept of a product life cycle, and observe that the shapes of such cycles have the general form of diagram 7.1.



The cycle is characterised by four stages. Following the introduction stage, when sales increase only slowly, the cycle enters the growth phase where a rapid increase in sales occurs. The cycle is convex over this section. However the cycle turns concave and levels out into the mature stage, before sales eventually decline. The length of time a product remains in any one stage will depend upon the individual characteristics of the product. The reasons given for the convex and then concave shape of the cycle is normally in terms of the diffusion of information between individuals. In the first period, one person buys the good and informs two other people of its advantages. In the second period these two consumers each tells two more, and this process generates a convex curve. However, after a time the economy runs out of uninformed people and the cycle flattens out.[2]

In this chapter, we provide an alternative explanation to the 'stylized-fact' of the shape of the product life-cycle, without making any assumptions about the diffusion of information. In fact we claim that even in a world of isolated individuals it is still possible to obtain a product life-cycle like the one in the diagram. The results are generated by the consumer's desire to gain information and then how he acts once the information has been received. We argue that learning affects consumer behaviour in two ways: firstly consumer behaviour is affected by the knowledge that the consumer has the ability to learn about unknown variables; secondly the accumulation of past knowledge affects the amount of information that the consumer has before him in the current period. Stoneman (1981) investigates how the interaction of Bayesian learning with an uncertain environment produces a sigmoid diffusion curve. He considers the problem of a firm deciding at each point in time the proportion of a new technology to adopt. He assumes "that the firm acts in a myopic manner", that is, does not look ahead to see that current decisions affect future parameters. Further he does not consider the case of a firm rejecting the new technology, on the basis of his experience showing that it fails to perform. He is able to generate a sigmoid diffusion curve for the individual firm, essentially by the process of the variance of returns being reduced through time. But the curve that is generated, is really the optimal plan of the firm conditional on all the information available in the current period. It is a series of optimal single period problems which the firm computes in the present, given the expected value of the observations that he expects to observe in the future. In this chapter we emphasise the importance of current decisions being made with reference to future utility. The consumer leans over time whether the product is good or bad. This in general does not produce a sigmoid curve for the individual but when behaviour is aggregated over individuals, it does for both products that in the long run are adopted, and those that are rejected.

In the next section of this chapter we look at a general solution to the optimal plan. In Section 3 we see how the solution to the optimal plan is modified by actual consumption patterns over time, in the light of information received. In the fourth section we look at the trend in demand elasticities along the life cycle. In Section 5, we aggregate the results of Section 3.

2. A General Solution to the Optimal Plan?

We can distinguish between two types of demand profiles: the consumer's planned profile which is worked out at the onset of the problem, and his actual consumption profile which shows his actual purchases of the good as he moves along in time and updates his expectations.

In the first case, the knowledge that he is able to learn about the unknown parameters in the problem for use in future periods, affects the consumer's current decisions. In the second case, the consumer's experience in the past affects his current decisions.

To solve the first problem, we set up a dynamic programming framework and the solution is found by backward induction. This was done in Chapter 4, for a two period environment, but now let the number of time periods, t, be t = 1, ... T. Then the problem in (4.9) can be restated as

$$\max_{x_t} EU_t = Eu_t + \frac{1}{1+i} EV_{t+1}. \qquad (7.1)$$

The consumer solves his optimal plan, by choosing the optimal decision variable in the final period, x_T^*. This optimal value will depend upon the purchases of the new good in the previous period and his experience from the previous period.

$$x_T^* = x_T^* (x_{T-1}, w_{T-1})$$

The consumer obtains the indirect function from these optimal values, and weights this maximum value function for every possible value of the random variable w_{T-1}. The consumer steps back one period, and solves his decision problem in T-1, taking account of the effect of x_{T-1} on the expected maximum value function in the final period. This backward induction process is carried on back to t = 1. At each stage, the optimal decision variable is specified as a function of the previous consumption level. Having solved the problem for x_1^*, the consumer will have a planned optimal consumption pattern, conditional upon all the information available at t = 1.

We wish to examine the optimal consumption plan of the consumer. Firstly we ask the question: is it possible to obtain an optimal consumption plan with a linear utility function? The short answer is, no. The solution is difficult because at each stage, the consumer needs to solve a maximisation problem and obtain an optimal solution to the decision variable. In practice it is not possible to obtain an analytical solution to this problem, simply because the utility function other than in the final period is extremely complex. And as we move further back to period 1, the utility function becomes progressively more complex. The problem can be solved numerically but the algorithm required to compute the backward induction process from t = T to t = 1 will be large. In chapter 4 we obtained numerical solutions for x_1^* in a two period model. In a T period model the numerical solution to x_{T-1}^* would be substituted back into the utility function and the expected indirect utility function could be computed for EVT-1 by integrating over the observations from T-2. This numerical process could be carried back to t=1. Alternatively a simple three period optimal plan can be solved by making the approximations outlined in Section 4. This allows us to obtain analytical solutions for both x_3^* and x_2^* which can then be used to solve the first period problem numerically.

3. Actual Consumption Patterns

The previous section considered the derivation of a consumer's optimal planned consumption profile. But there is no reason why the consumer should necessarily stick to this plan as he moves along. His optimal consumption plan will vary at each point in time. Thus at the beginning of the first period, having worked out his optimal plan, the consumer purchases the optimal amount of the new good in the first period. Having carried out this purchasing decision, the consumer observes the results and updates the parameters accordingly.

With his set of updated parameters he moves into the second phase of his consumption plan. He then purchases the optimal amount of the new good, observes the realisation of the random variable, updates the parameters, and moves into the third period. He carries out this process for $t = 1, \dots T$. At each period he has an optimal consumption plan, though he only actually purchases the initial optimal value from any particular plan. By looking at this stream of initial purchases from each plan, we will see how the consumer actually behaves. To calculate his actual behaviour, we need to make an assumption about the experience that the consumer observes; and in order to simplify the problem we need to reduce the length of the optimal plan. Suppose that, although we allow the consumer to make decisions over a number of time periods, in each period he is only allowed to look forward one period. That is, the optimal plan consists of optimal decision variables, given current information over only two periods. In the current period, the consumer knows that he will be making decisions over a number of periods in the future, but he only considers the impact of learning and the value of information on utility in the next period.

To obtain the optimal plan the consumer maximises the flexible budget constraint of the linear utility function discussed in Chapter 4. His optimal plan changes over time as new information becomes available. How does this new information affect the optimal plan? Suppose the consumer has solved the optimal two-period plan in equation (7.1) for $t = 1$, and purchased a quantity of the new good x_1^*. He then observes the realisation of the random variable $\bar{\varepsilon}_1$. We impose a time subscript on this experience, to show it is the observed sample statistic from the first period. Suppose $\bar{\varepsilon}_1 = \mu_o - \eta$. Then the consumer updates his subjective beliefs to obtain a new estimate of the mean, that he will use as his prior in the next decision problem.

$$\mu_1 = \frac{\mu_o + (\mu_o - \eta)\, x_1^*}{x_1^* + \phi}$$

$$\sigma_1^2 = \frac{1}{\Phi_1} = \frac{1}{\phi + x_1^*} \tag{7.2}$$

where x_1^* denotes the quantity of the optimal purchases of the new good in the first period.

Armed with these new beliefs in (7.2), the consumer moves into the second period and faces the following problem:

$$\max_{x_2} \ EU_2 = \mu_1 x_2 + \frac{1}{1+i} EV_3 \tag{7.3}$$

where EV_3 is the expected maximum value function for the third period. It is computed by assuming that the consumer only looks one period ahead, and at the beginning of the second period acts as if the third period is the final period.

$$EV_3 = (1 + r)\{\bar{M} - px_2\} \int_{-\infty}^{\bar{\varepsilon}_2^*} f(\bar{\varepsilon}_2)d\bar{\varepsilon}_2 + \frac{1}{p} \int_{\bar{\varepsilon}_2^*}^{\infty} \mu_2 f(\bar{\varepsilon}_2)d\bar{\varepsilon}_2 \tag{7.4}$$

where $\mu_2 = \dfrac{\mu_1 \Phi_1 + \bar{\varepsilon}_2 x_2}{\Phi_1 + x_2}$

and $\bar{\varepsilon}_2 \sim N\left(\mu_1, \dfrac{1}{\Phi_1} + \dfrac{1}{x_2}\right)$

and μ_1 and Φ_1 are defined by equation (7.2).

The income variable in (7.4) is different from the income variable specified earlier. Previously income was given to the consumer at the beginning of the problem, and the consumer, in the flexible budget case was allowed to spread this income over two or three periods. We still wish to retain the flexible nature of the budget constraint, but do not wish to make it too flexible. For example if consumption today affects the budget 'n' periods hence, where n is a large number, then the solution to the problem is difficult; since we have to then specify an 'n' period budget constraint. Instead we suppose that the consumer receives a fixed income stream M, every period , but is allowed to allocate this income over the current period and the next period. But of course at any time the consumer not only has income M but also accumulated savings from the previous periods. The budget constraint can be written as

$$y_{t+1} + px_{t+1} = M_{t+1} + (1+r)\{M_t + B_t - (y_t + px_t)\} \qquad (7.5)$$

and $$B_t = \sum_{\tau=1}^{t-1} (M_\tau - px_\tau)(1+r)^{t-\tau}$$

We are making a similar assumption with respect to the length of the budget constraint as we did in respect of the optimal plan. There, the consumer was allowed to look only one period ahead even though he was aware that there were many periods in front of him. Here again, the consumer recognises that savings today can be used tomorrow, but not that savings today can be used further ahead, even though the consumer observes that today's income includes accumulated savings over many past time periods. If income streams are constant: $M_t = M_{t+1} = M_{t+i} = M$, then let

$$\overline{M} = \{M\left(\frac{2+r}{1+r}\right) + B\} \qquad (7.6)$$

and the budget constraint can be written

$$y_{t+1} = (1+r)\{\overline{M} - px_t\} - px_{t+1} \qquad (7.7)$$

The consumer now maximises (7.3) and obtains an optimal value for x_2, which consequently represents his actual purchases of the new good in the second period. The consumer will again observe the realisation of his experience, this time from the second period. Suppose again,

$\bar{\varepsilon}_2 = \mu_o - \eta$

$$\therefore \qquad \mu_2 = \frac{\mu_1 \phi_1 + (\mu_o - \eta)\, x_2^*}{\phi_1 + x_2^*}$$

Substituting from (7.2):

$$\mu_2 = \frac{\mu_o - \eta \sum_{t=1}^{2} x_t^*}{\phi + \sum_{t=1}^{2} x_t^*} \qquad (7.8)$$

and $\phi_2 = \phi_1 + x_2^* = \phi + \sum_{t=1}^{2} x_t^*$

The consumer now moves into the third period armed with his new beliefs: (μ_2, ϕ_2). This continual updating process, adapting the optimal pla in each period, gives the consumer's actual purchases in each period: $x_1^*, x_2^*, \ldots x_T^*$. We wish to observe this profile.

In general at any point in time τ, the consumer will be looking ahead one period into $\tau + 1$. He must choose a value of x_τ to maximise his expected utility over the two periods, τ and $\tau + 1$, given the updated values of the parameters, which he has been learning about over the last $t = 1, \ldots \tau - 1$ periods. Again, the rate of interest and the rate of discount are set equal to one. The consumers problem is:

$$\max_{x_\tau} \; EU_\tau = \mu_{\tau-1}x_\tau \; + \; EV_{\tau+1} \tag{7.9}$$

$$\text{where } EV_{\tau+1} = \bar{M} - px_\tau + \frac{(\bar{M}-px_\tau)}{p}\left\{(\mu_o-p) - \frac{\eta\,\Omega}{\Phi+\Omega}\right\}\int_{m^*}^{\infty} f(m)dm$$

$$+ \frac{(\bar{M}-px_\tau)}{p}\left(\frac{x_\tau}{\Phi+\Omega}\right)^{\frac{1}{2}} \cdot \frac{1}{(x_\tau+\Phi+\Omega)^{\frac{1}{2}}} \cdot \frac{e^{-\left(\frac{m^*}{2}\right)^2}}{\sqrt{2\pi}}$$

$$\text{and } m^* = \left\{(p-\mu_o) + \frac{\eta\,\Omega}{(\Phi+\Omega)}\right\}(x_\tau+\Phi+\Omega)^{\frac{1}{2}}\left(\frac{\Phi+\Omega}{x_\tau}\right)^{\frac{1}{2}}$$

$$\text{and } \mu_{\tau-1} = \mu_o - \frac{\eta\,\Omega}{(\Phi+\Omega)}$$

$$\text{and } \Omega = \sum_{t=1}^{\tau-1} x_t^* .$$

It can be seen that over time, as the consumer works out a new two period optimal plan, his prior beliefs for each successive plan have changed due to updating the mean and an increased degree of precision. By imposing initial values on the parameters μ_o, Φ, M and p, and specifying the discrepancy between the consumer's initial beliefs and his experience η, we can solve equation (7.9) numerically for

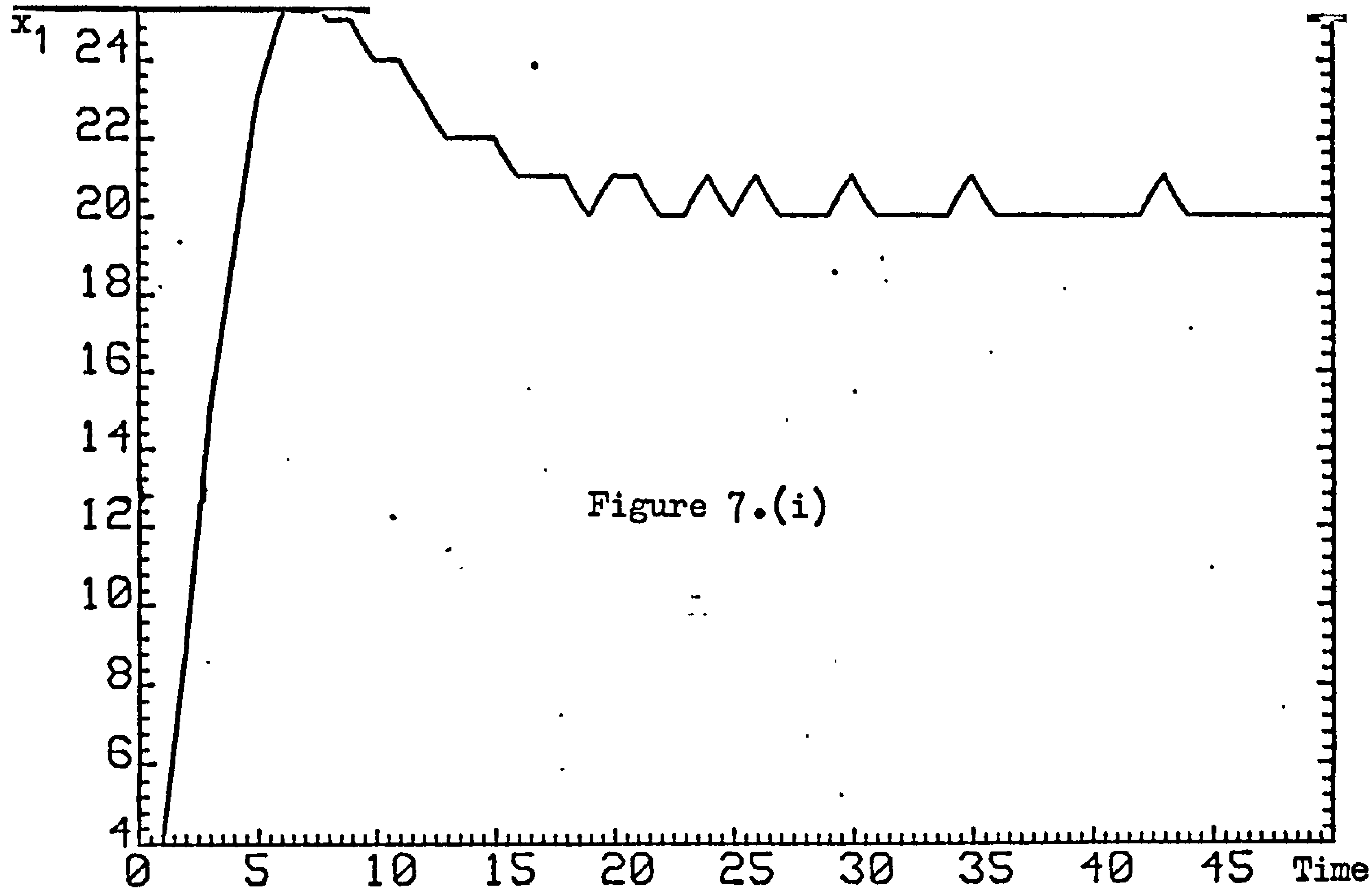


Figure 7.(i)

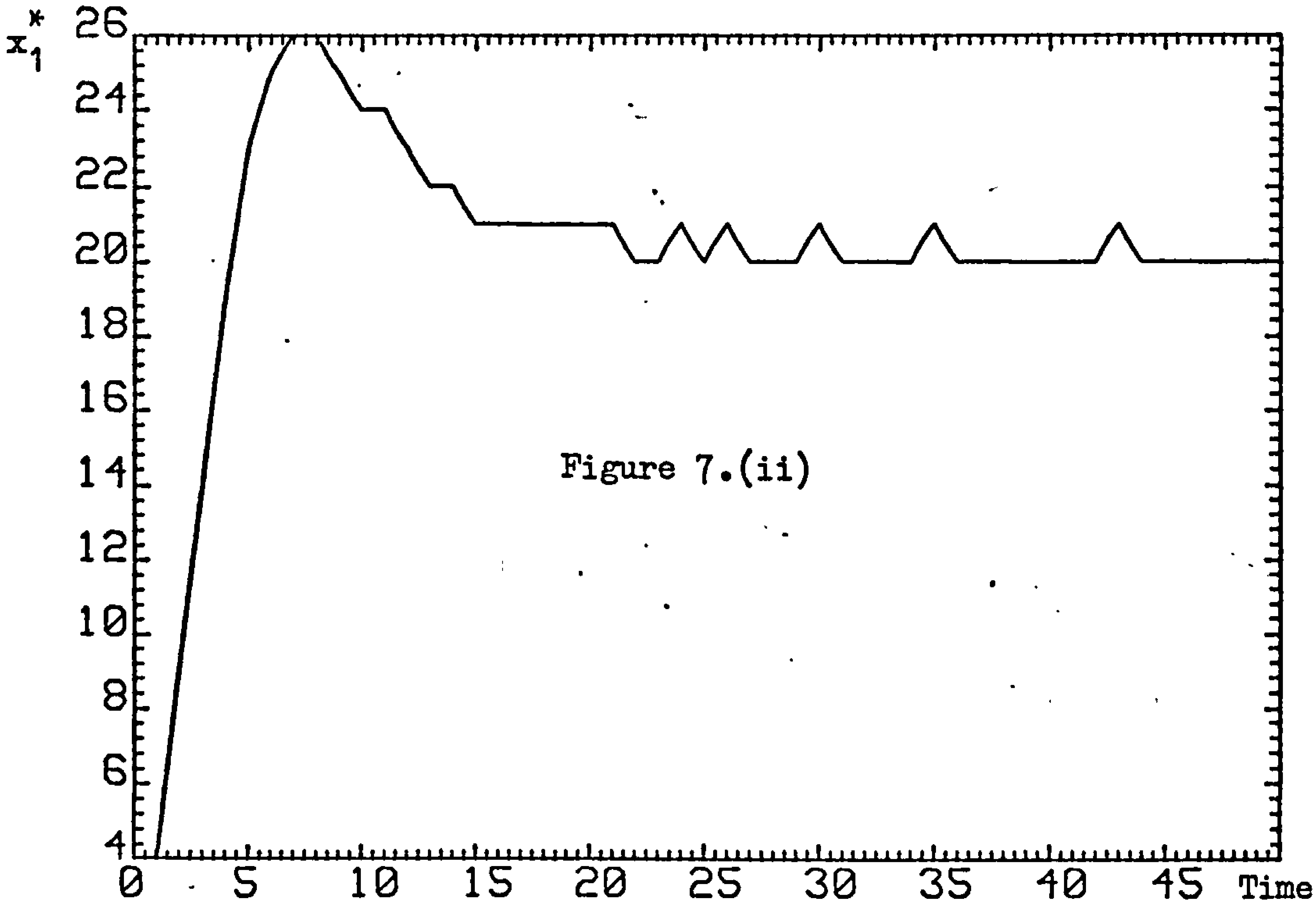


Figure 7.(ii)

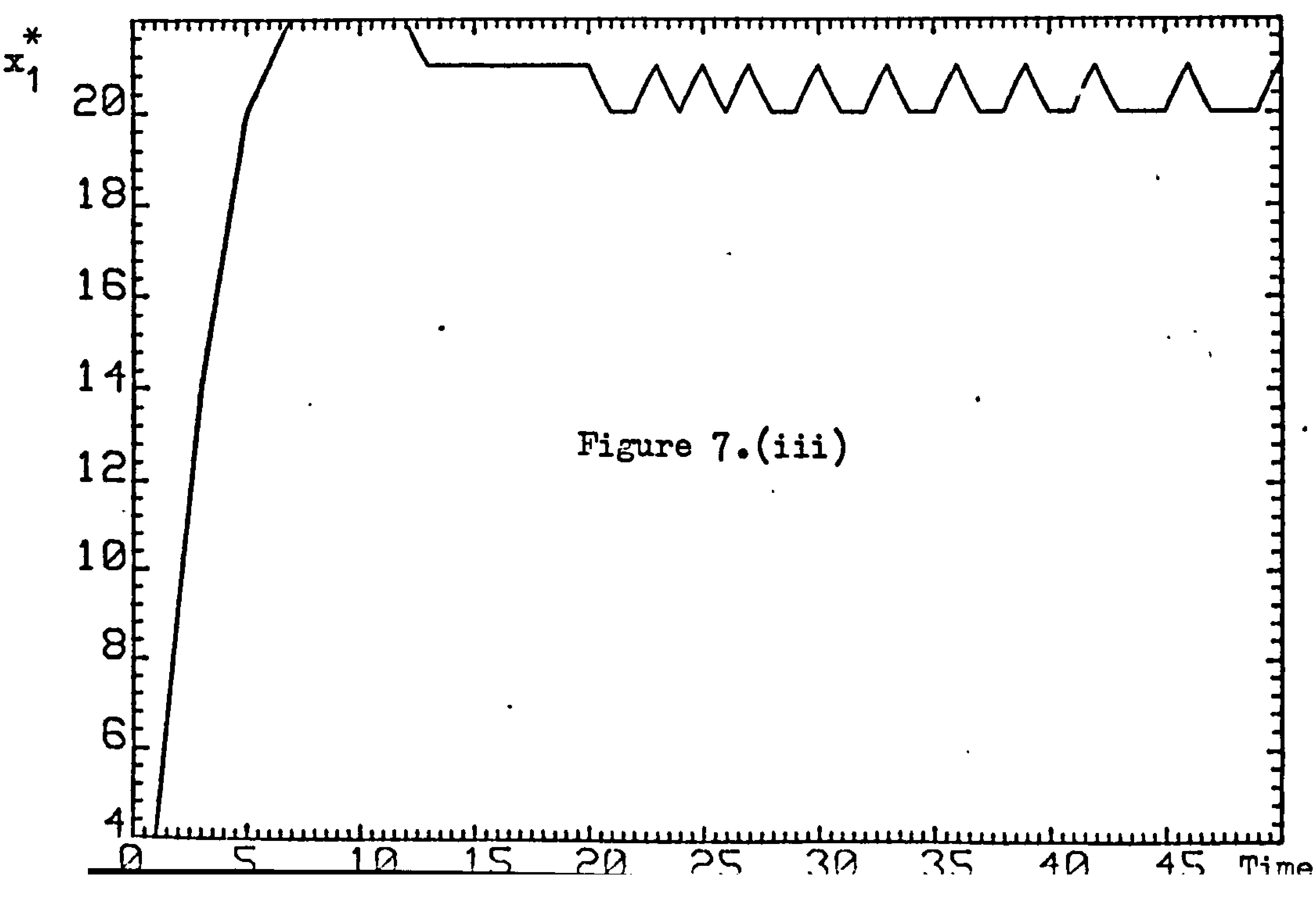


Figure 7.(iii)

each time period $t = 1, \ldots t$, to obtain the time profile of optimal purchases.

We now carry out a series of comparative statics exercises to see the effect of changing the parameter values on the pattern of purchases. The most important parameters values are the relative values of μ_o, p and η. The difference between μ_o and p reflects the difference between initial expectations of product quality and price. The value of $\mu_o - \eta$ represents the long run expectation of product quality, based on the consumer's experience. For a linear utility function, whether the consumer adopts the new product or not in the long run depends upon whether $\mu_o - \eta \gtrless p$. The long run equilibrium demand, is either to purchase only the new good or to purchase none of it. In examining the effect of parameter values on the solution to the model, we are interested in the speed of adjustment to the final equilibrium. The tables and figures below show the effects of parameter changes.

In table 7.1 we compare the effect of a difference in initial expectations on the pattern of purchases. It can be seen that row (i) has $\mu_o > p$, and row (ii) has $\mu_o < p$, which with a linear utility function in a temporally independent environment would result in the consumer spending all his budget on the new good in (i) and none in (ii). In this case though, the small difference does not affect the pattern of purchases. This is illustrated in figures 7.(i) and 7.(ii). Comparing row (ii) with (iii), the greater is long run expectations of product quality, the quicker the consumer

Table 7.1

	μ_o	p	ϕ	η	M	max x_1^*	time to reach steady state
(i)	0.51	0.5	1.0	-0.01	10	26	19
(ii)	0.49	0.5	1.0	-0.03	10	26	22
(iii)	0.49	0.5	1.0	-0.012	10	22	22

adjusts to his long run position. Figure 7.(iii) shows a more gradual movement. It can be seen from the figures that long run equilibrium exhibits a series of blips; these occur because of the discrete nature of the decision variables x_t^* is only allowed to take on integer values. Of more concern are the much larger blips that occur when $p < \mu_o - \eta$. These are illustrated in figures iv - vi and table 7.2.

Table 7.2

	μ_o	p	ϕ	η	M	max x_1^*	time to reach steady state
(iv)	0.49	0.5	0.01	0.001	10	20	20
(v)	0.48	0.5	0.01	0.001	10	18	12
(vi)	0.47	0.5	0.01	0.001	10	16	10

These oscillations are due to income effects as a result of the flexible budget constraint. It was shown to be optimal not to consume any of the old good, since the consumer can always delay this decision. The result is that the consumer accumulates savings over time, and if he purchases none of the old good either, then all

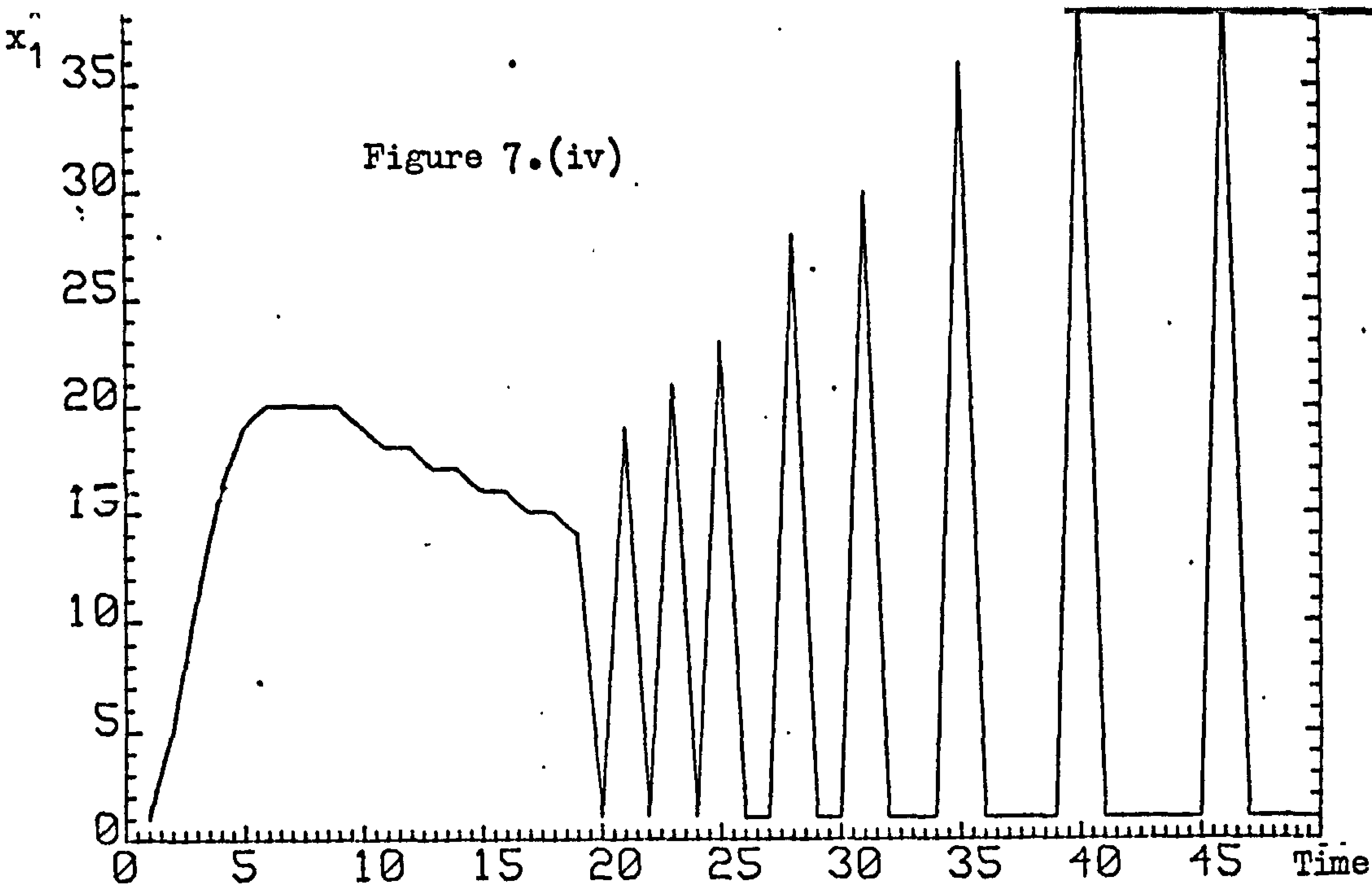


Figure 7.(iv)

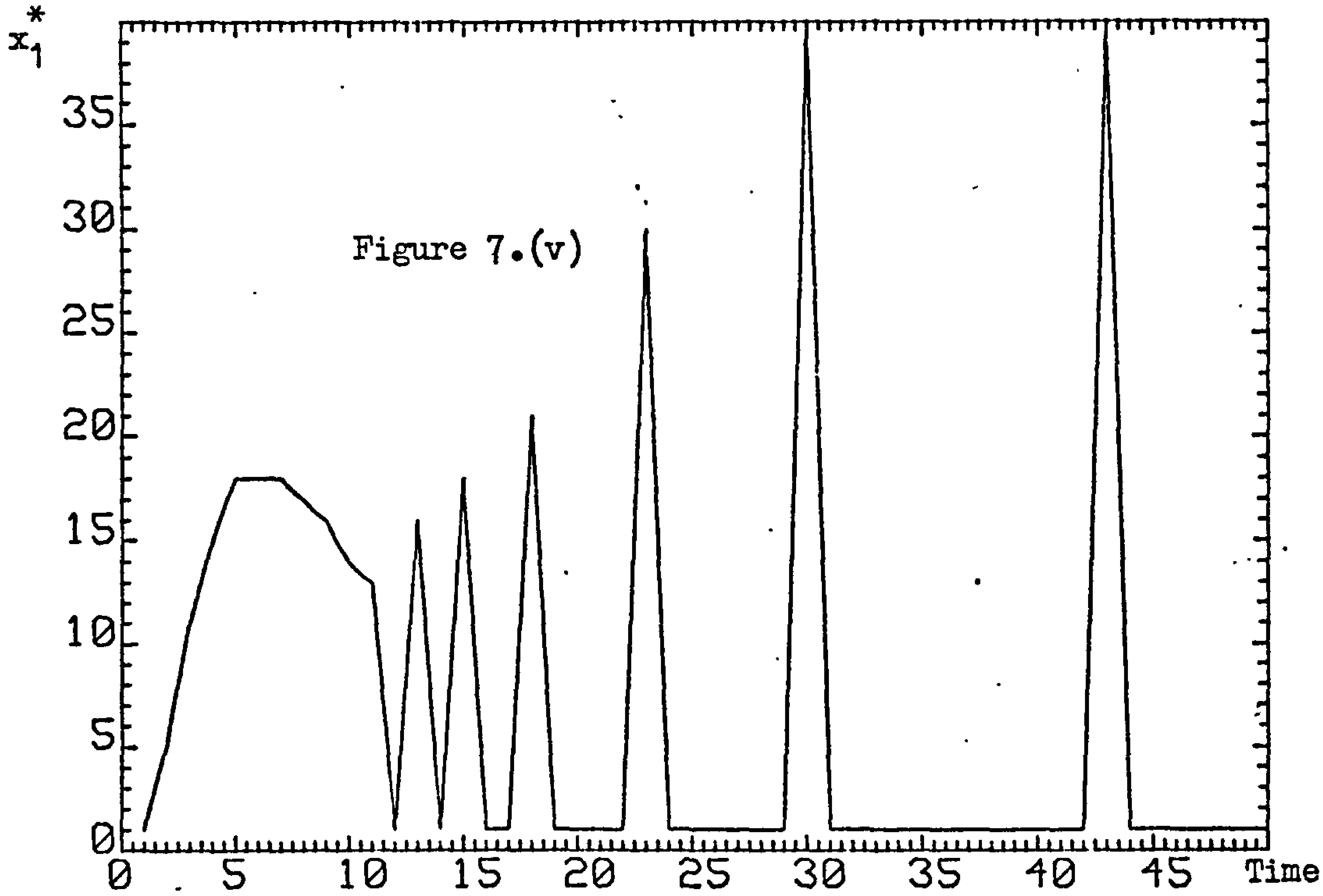


Figure 7.(v)

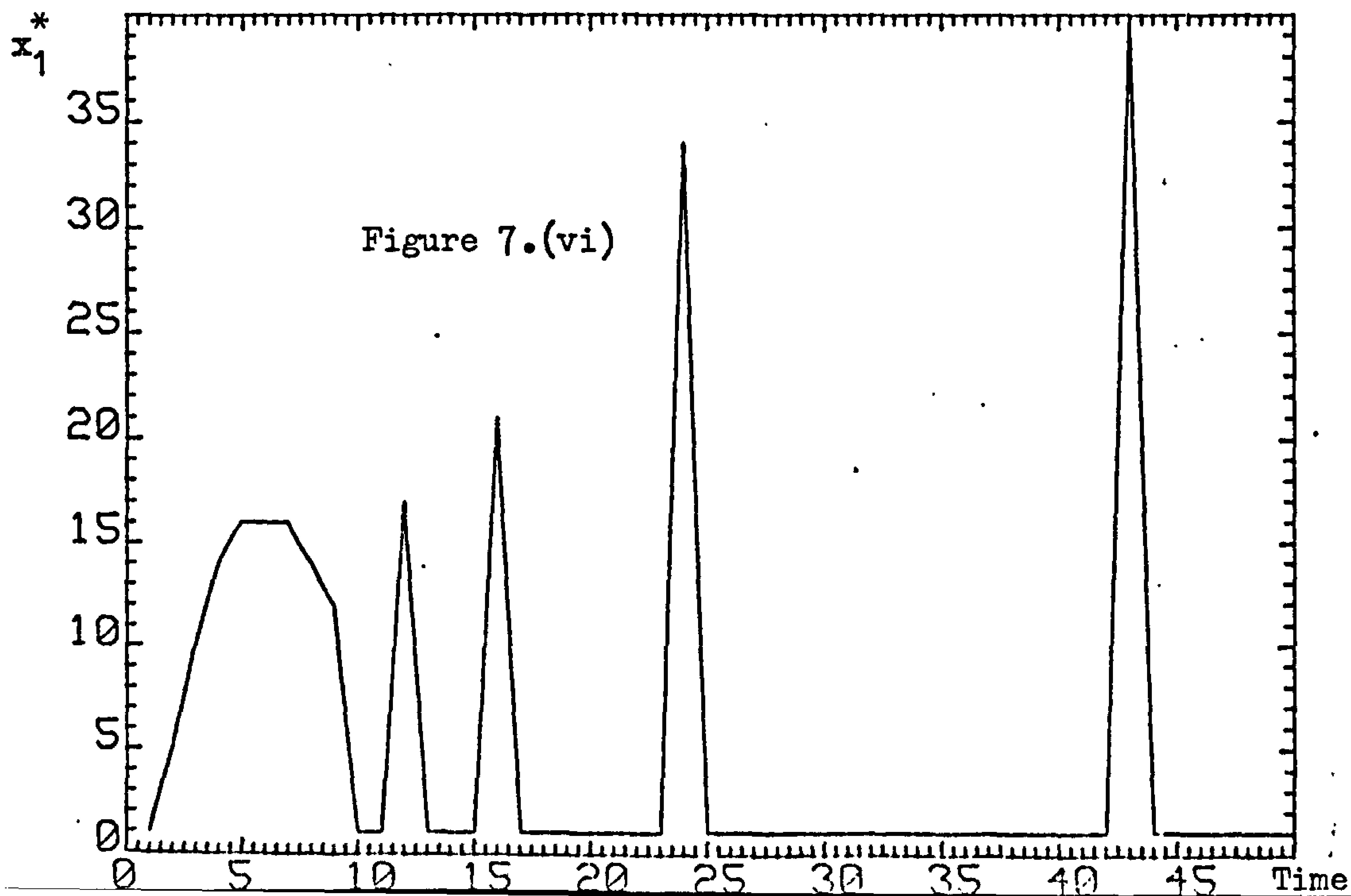


Figure 7.(vi)

his income is saved. The consumer becomes very rich. He is so rich that he is prepared to sample once again, even though he is extremely confident that product quality is less than the price. But because he does not know this for certain, it is worthwhile him sampling again. This explains why the oscillations get further apart through time, since to keep on sampling, he must keep accumulating more and more income. Also the diagrams show that as the difference in quality and price diverge, these oscillations occur less frequently. We can also see that the initial sampling period decreases, as the difference between p and μ_o increases. In order to overcome the problem of these oscillations appearing in the remaining comparative static exercises, we impose the restriction that if ever $x_t^* = 1$, and $t > 1$, then $x_t^* = 1 \forall \tau$, $\tau = t, \ldots T$.

Table 7.3

	μ_o	p	ϕ	n	M	max x^*	time to reach steady state
(vii)	0.51	0.5	1.0	0.02	10	19	21
(viii)	0.49	0.5	1.0	0.0	10	19	21
(ix)	0.49	0.5	1.0	-0.005	10	20	39

The first two rows in table 7.3 show the same long run expectation of product quality as 0.49. But the former has a higher initial expectation. This difference does not have a great impact on the profile of the purchaser. Both cases in figures vii and viii look very similar. Case ix shows the profile of a good whose initial and long run quality expectations are less than the price, but whose long run expectations are less pessimistic than the initial ones.

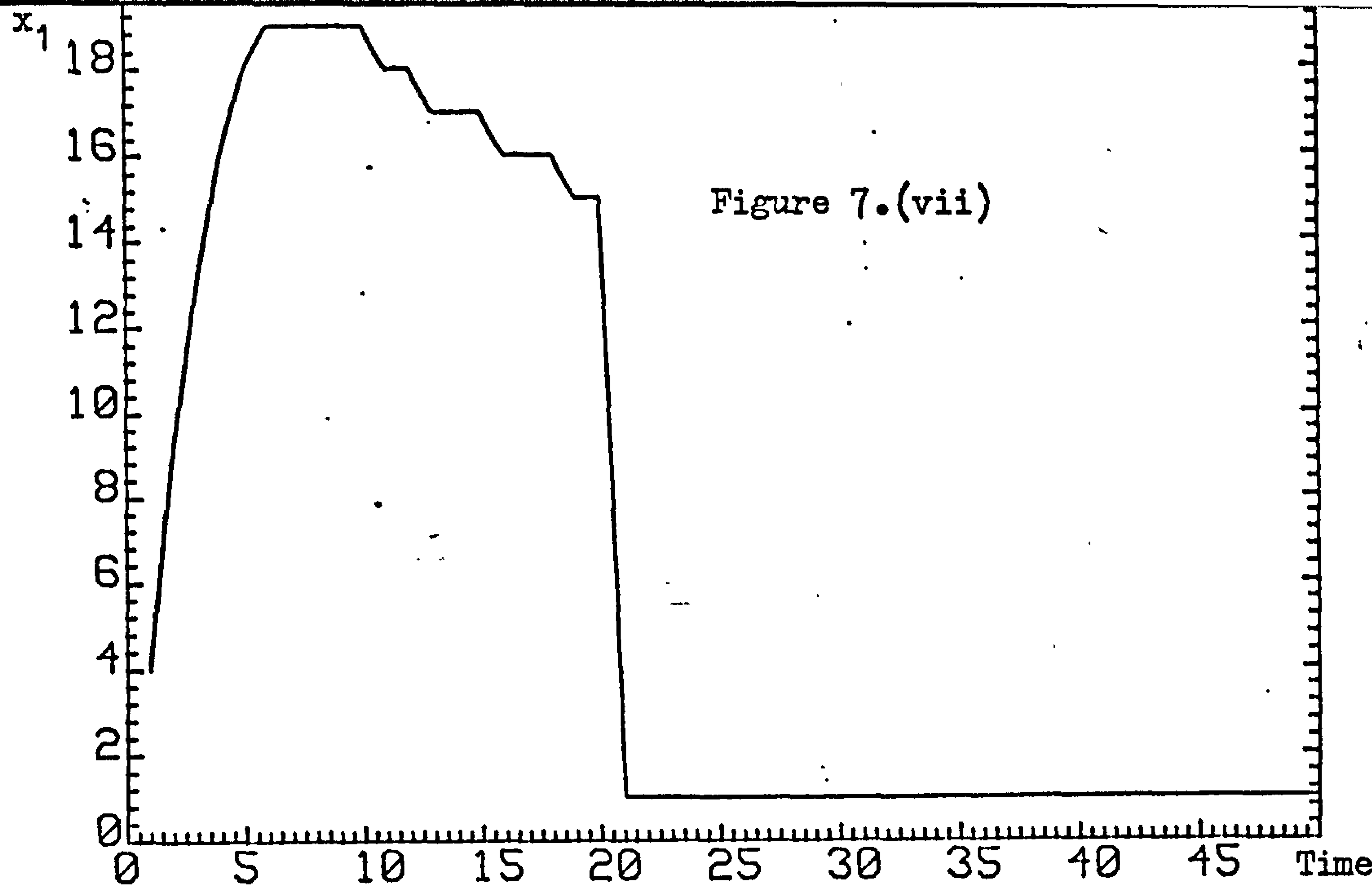


Figure 7.(vii)

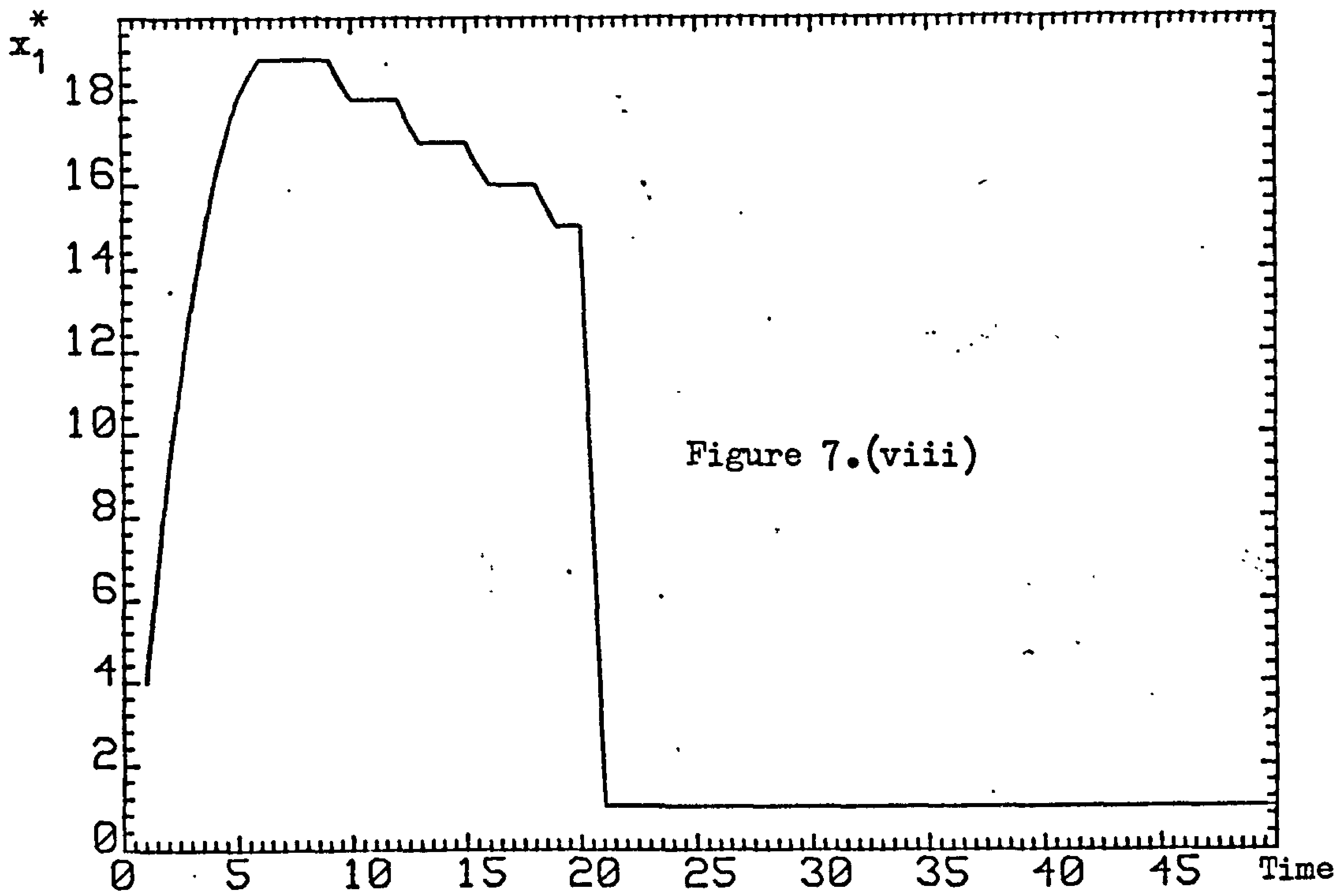


Figure 7.(viii)

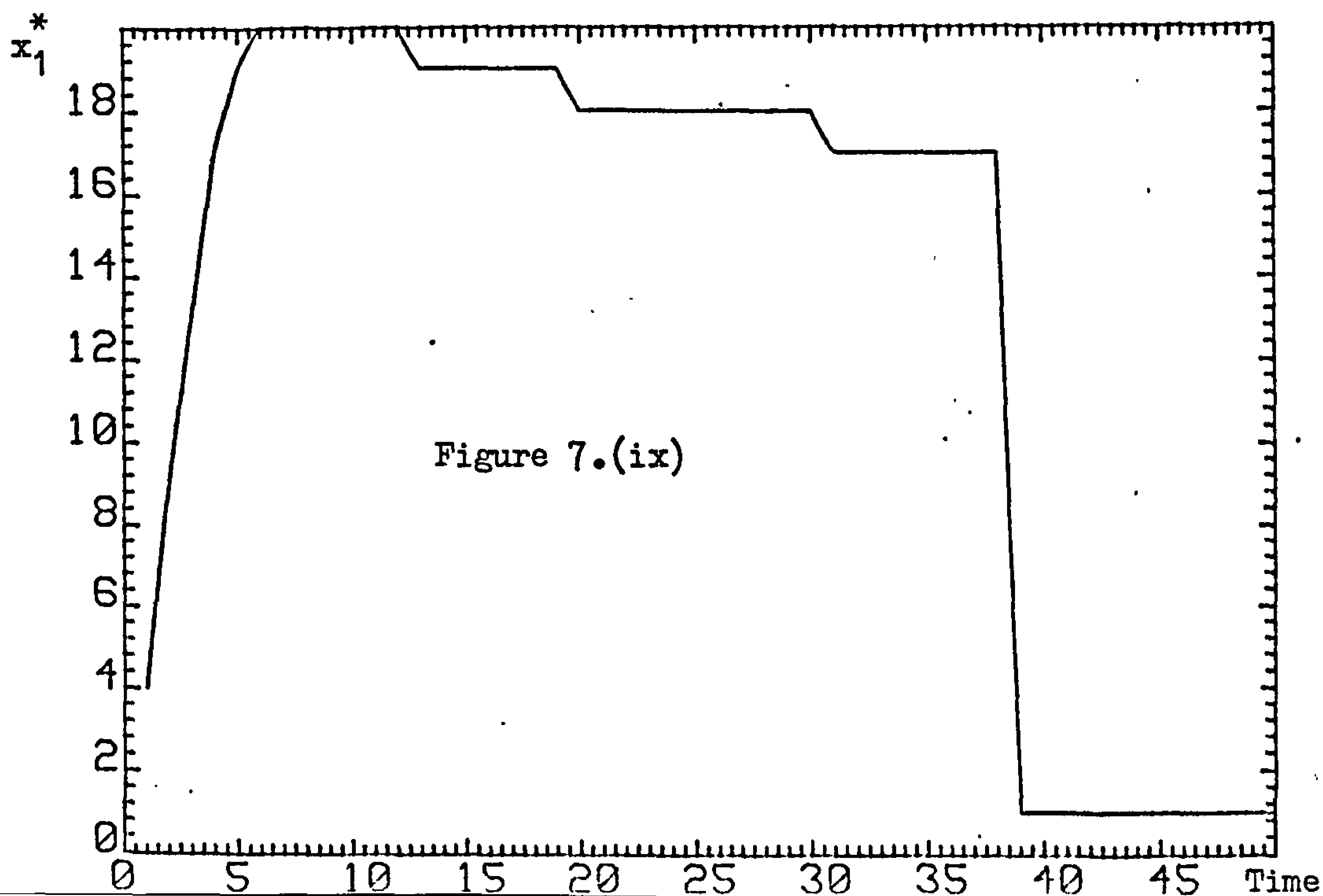


Figure 7.(ix)

Again the consumer ends up not purchasing the good, but takes a long time to reach the long run equilibrium.

Table 7.4

	μ_0	p	ϕ	η	M	max x_t^*	time to reach steady state
(x)	0.49	0.5	0.01	0.01	10	18	13
(xi)	0.49	0.5	0.01	0.1	10	10	5

Both case (x) and (xi) have $\mu_o < p$ and $\mu_o - \eta < p$. In the long run the consumer buys neither, but he initially buys more of case (x) and takes longer to decide that its actual quality is less than its price. In both table 7.4 and (viii) and (ix) from 7.3, although initial expectations are less than price, and this is confirmed by his experience, the consumer continues to increase his purchases of the good. This can be explained by proposition 4.5, which states that as the degree of precision increases, at low levels of precision the optimal quantity of purchases increase. As the consumer accumulates past purchases, this acts as an increase in the initial degree of precision.

Table 7.5

	μ_o	p	ϕ	η	M	max x^*	time to reach steady state
(xii)	0.52	0.5	1.0	0.03	10	19	21
(xiii)	0.52	0.5	1.0	0.01	10	24	

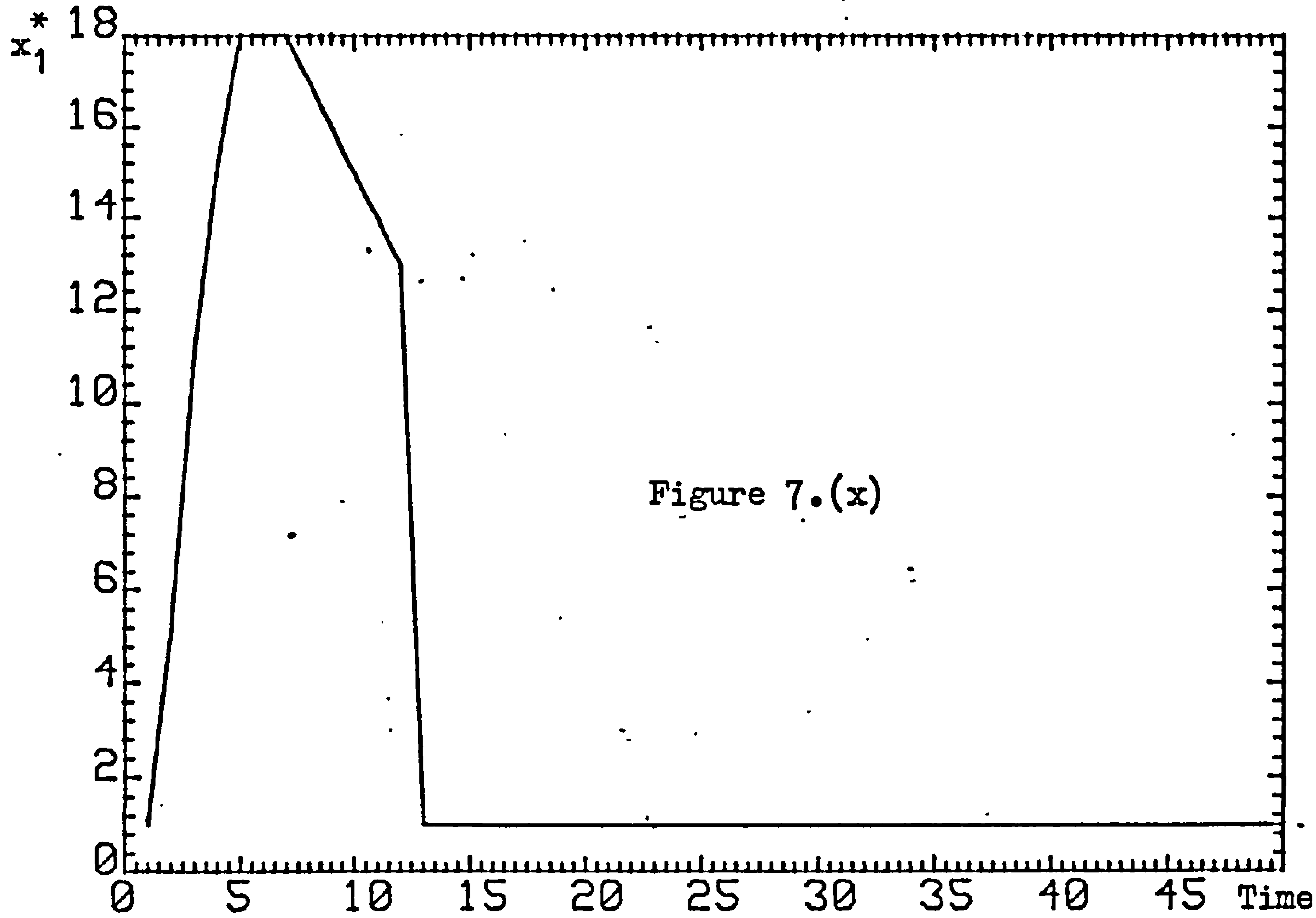


Figure 7.(x)

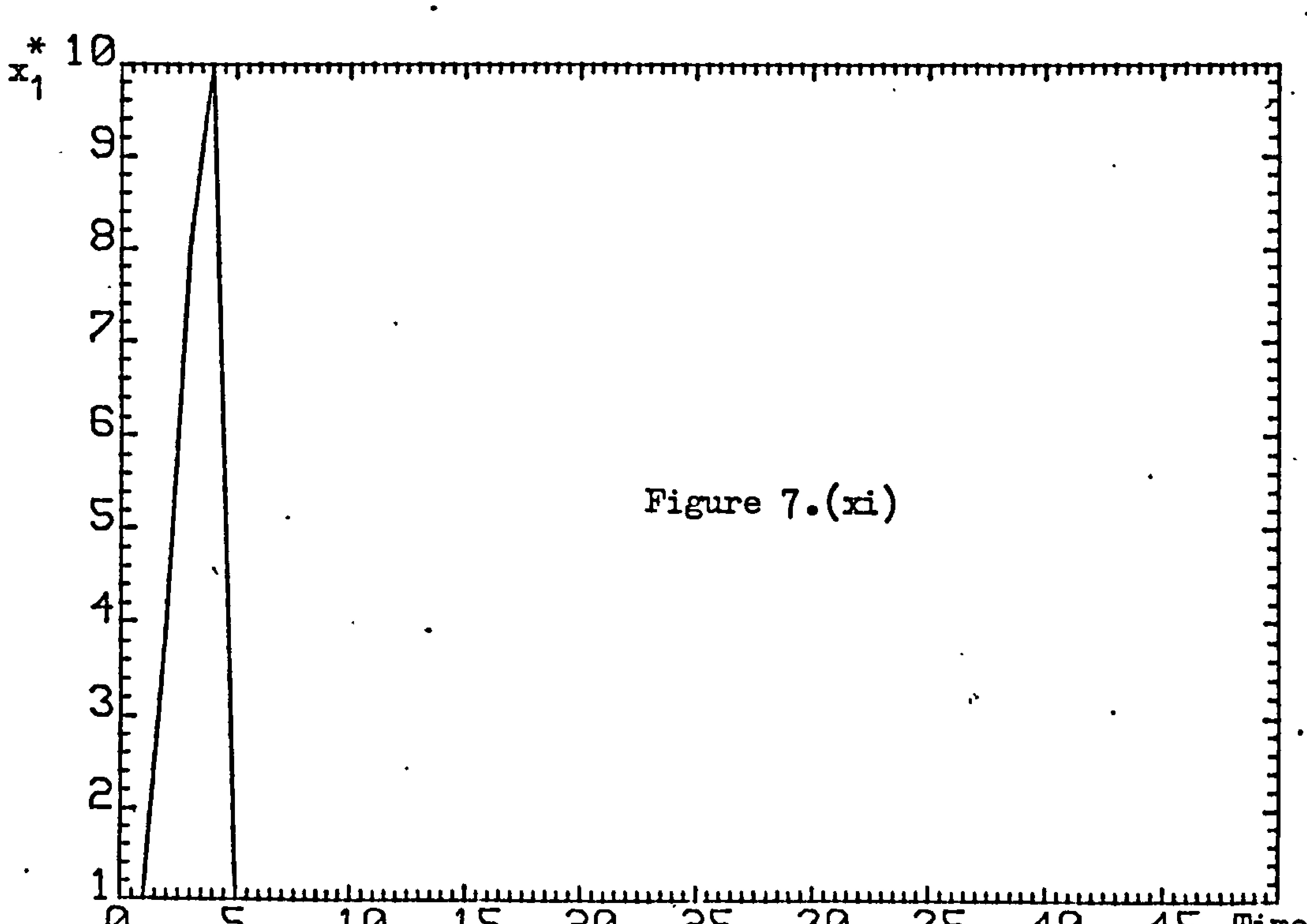
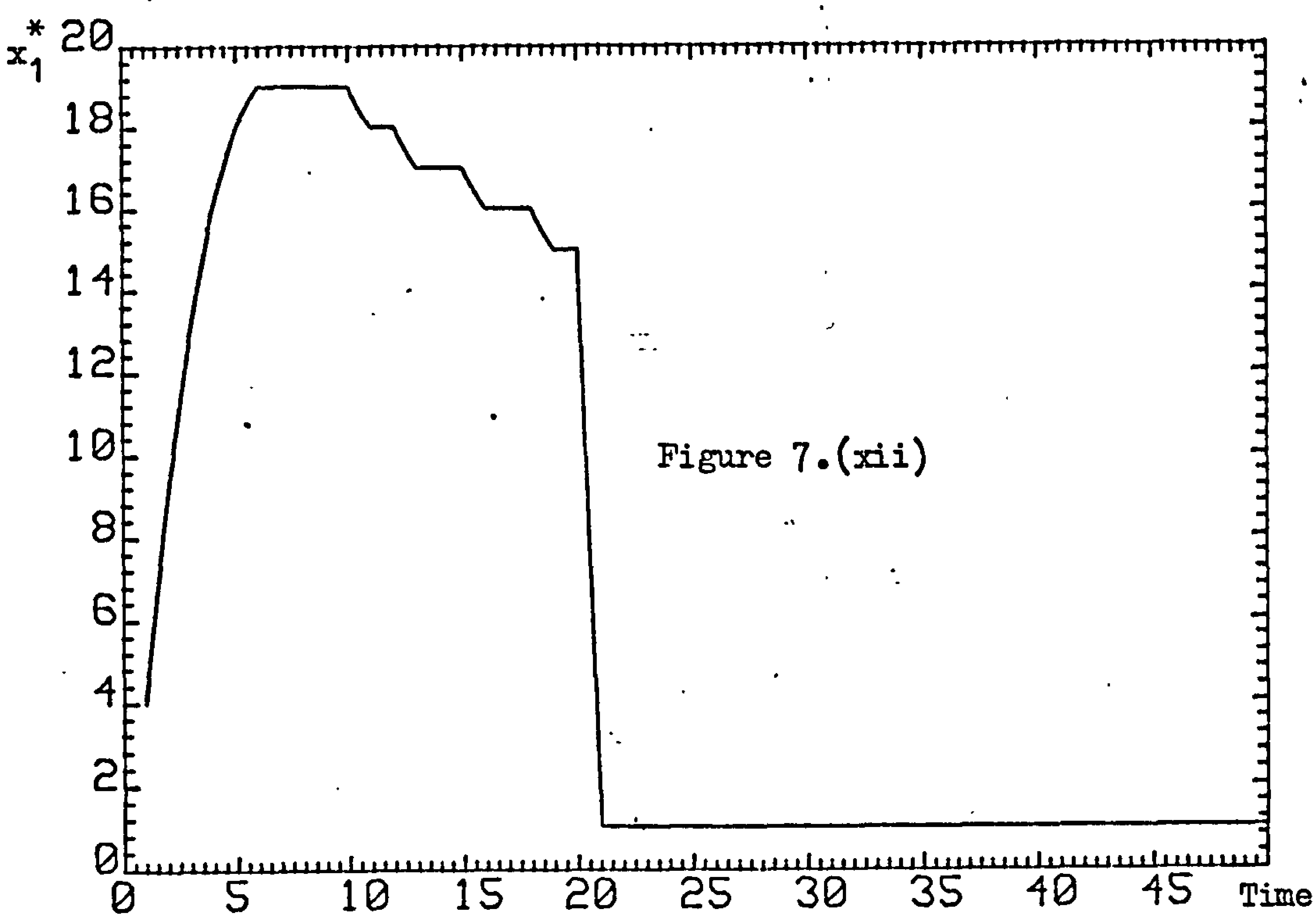


Figure 7.(xi)


x_1^*
Figure 7.(xii)
Time

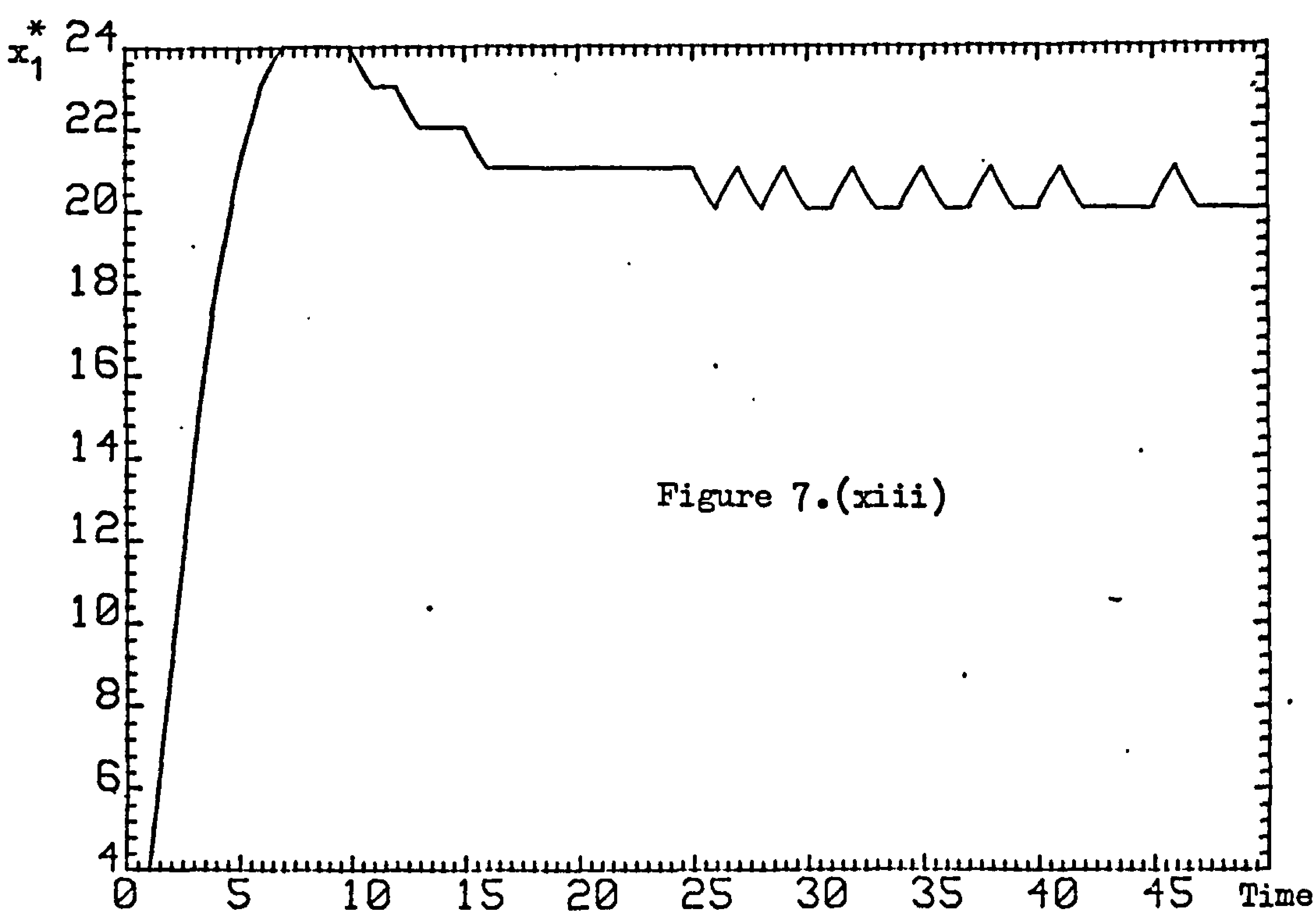

x_1^*
Figure 7.(xiii)
Time

Table 7.5 has the initial quality expectation greater than price, but long run expectations fall in both cases. In the former they fall below price and in the latter they stay above it.

4. Elasticities along the Life Cycle

The life cycle shows the profile of optimum demands for the good. Each point on the cycle is taken from an associated demand curve in each time period. In this section we look at the shape of these implicit demand curves, and how they change over time.

Consider the linear utility function and the flexible budget constraint in Chapter 4. From equation (4.32), we observe that if $p - \mu_o$ is small, and ϕ is small, then $m^* \simeq 0$. This implies:

$$\frac{e^{-\left(\frac{m^*}{2}\right)^2}}{\sqrt{2\pi}} \simeq \frac{1}{\sqrt{2\pi}} \quad \text{and} \quad \int_{m^*}^{\infty} f(m)dm \simeq \frac{1}{2}$$

Differentiating equation (4.30) w.r.t. x_1 using (4.32) and proposition 4.9 :

$$\frac{dEU_1}{dx_1} = \mu_o + \frac{1+r}{1+i} \cdot \frac{(M-px_1)}{2p} \cdot \frac{1}{(\phi + x_1)^{3/2}} \cdot \left(\frac{\phi}{x_1}\right)^{\frac{1}{2}} \cdot \int_{m^*}^{\infty} mf(m)dm$$

$$- \frac{1+r}{1+i}\left[\int_{m^*}^{\infty} \left\{ \mu_o + \left(\frac{x_1}{\phi}\right)^{\frac{1}{2}} \cdot \frac{m}{(\phi + x_1)^{\frac{1}{2}}} \right\} f(m)dm + p \int_{-\infty}^{m^*} f(m)dm\right] \tag{7.10}$$

Setting this expression equal to zero and making use of the approximations, we can simplify (7.10):

$$\frac{(M-px_1)}{2p} \cdot \frac{\phi}{x_1} \approx (x_1 + \phi)$$

$$\therefore \quad x_1^2 + \frac{3}{2}\phi - \frac{M}{2p} (1+r)\phi = 0 \qquad (7.11)$$

Solving the quadratic and imposing the restriction that $x_1^* > 0$:

$$x_1^* \approx -\frac{3}{4}\phi + \frac{1}{4}\sqrt{9\phi^2 + \frac{8\phi}{p} M(1+r)} \qquad (7.12)$$

This value of x_1^* can be substituting into (4.30) to obtain the indirect utility function V_1.

Elasticity of demand χ is defined as $\frac{dx_1^*}{dp} \cdot \frac{p}{x_1^*}$, and from (7.12) we obtain $\frac{dx_1}{dp}$;

$$\frac{dx_1^*}{dp} \approx -\phi\frac{M(1+r)}{p^2} \cdot \frac{1}{\left[9\phi^2 + \frac{8\phi}{p} M(1+r)\right]^{\frac{1}{2}}}$$

$$\therefore \chi \approx \frac{-4\phi M(1+r)}{p\left[9\phi^2 + \frac{8\phi}{p} M(1+r) - 3\phi\left\{9\phi^2 + \frac{8\phi}{p} M(1+r)\right\}^{\frac{1}{2}}\right]} \qquad (7.13)$$

We are interested in how this elasticity changes over time. If we make the further assumption that $\eta = 0$, that is, that the experience the consumer observes is exactly what he expected, then the only parameter that changes over time is the degree of precision, which is affected by cumulated purchases. Thus the movement of elasticity along the time profile is summarised by $\frac{d\chi}{d\phi}$.

Proposition 7. If elasticity is measured by equation (7.13), then over the life cycle elasticity increases (more negative).

Proof Differentiate equation (7.13) w.r.t. ϕ

$$\frac{d\chi}{d\phi} = \frac{\frac{6}{p}\phi^2 M(1+r)\{\ 6 - \left[18\phi + \frac{8}{p} M\,(1+r)\right]\left[9\phi^2 + \frac{8\phi}{p} M\,(1+r)\right]^{-\frac{1}{2}}\}}{[\text{ denominator } (7.13)\,]^2}$$

But [denominator 7.13$]^2 > 0$, and it can easily be shown that the numerator of the above expression is always negative

$\therefore \quad \frac{d\chi}{d\phi} < 0$: elasticity becomes more negative

Q. E. D.

So that along the life cycle the demand for the product becomes more elastic. The intuition behind this result is that in the early periods, information gain is high, and the consumer has an inelastic demand for sampling purposes. However over time, information acquisition means that the marginal value of information declines. Thus the consumer no longer demands the good for sampling reasons, and consequently the demand becomes more sensitive to price changes.

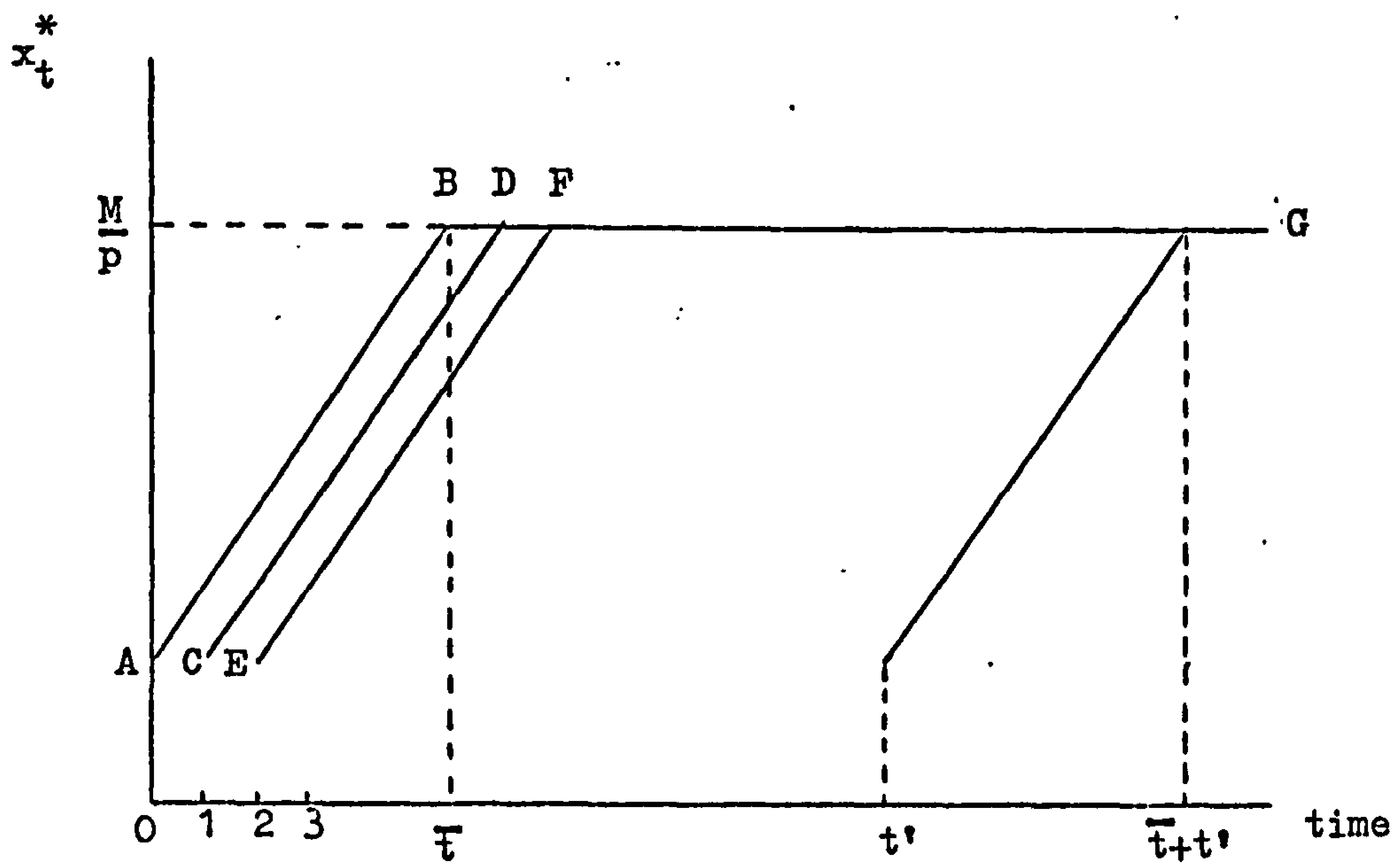
5. Aggregate Consumption Patterns

The previous section considered the pattern of an individual consumer's purchases once he was aware that the new good existed. Let us suppose that all the consumers in the economy are identical but that they become aware of the existence of this new good at different points in time. Specifically, in each period a different consumer buys the good for the first time. If all consumers are the same, except for this knowledge of the existence of the new product, then the time profile of purchases for each individual consumer will be identical; but each profile will start one period later.

The aggregate time profile will be the summation of the individuals' purchases. At any point in time the aggregate purchases of the new good will be made up of consumers at different points on their individual time profiles. Consumers will be at different points on the learning process.

What will be the shape of the aggregate time profile? Suppose that the continual two period optimal plan of the previous section results in purchases of the new good being a linear function of the time period, up to the point where the budget is exhausted by purchases of the new good. If the product is of a high quality relative to its price, so that as a result of the consumer's experience, the long run equilibrium is to spend all his budget on the new good, then

$$\lim_{t \to T} x_t^* = \frac{M}{p}$$



In the diagram, ABG is the time profile of purchases for an individual who purchases the good for the first time in $t = 0$. CDG is the time profile of an identical individual who becomes aware of the good one period later. Each new consumer who enters the market will have a similar purchasing profile. To find the aggregate profile, we sum the individual profiles vertically.

If we write the first individual's profile as:

$$x = a + bt \qquad 0 \le t \le \bar{t}$$
$$x = \frac{M}{p} \qquad t > \bar{t}$$

The profile for the second individual will be:

$$x = (a - b) + bt \qquad 1 \le t \le \bar{t} + 1$$
$$x = \frac{M}{p} \qquad t > t + 1$$

and total sales at $t = 1$ are

$$X_1 = 2a + b$$

The third consumer has profile:

$$x = (a - 2b) + bt \qquad 2 \le t \le \bar{t} + 2$$
$$x = \frac{M}{p} \qquad t > \bar{t} + 2$$

and total sales at $t = 2$ are

$$X_2 = 3a + 3b.$$

For the fourth consumer:

$$x = (a - 3b) + bt \qquad 3 \leq t \leq \bar{t} + 3$$
$$x = \frac{M}{p} \qquad t \nleq \bar{t} + 3$$

and total sales at t = 3 is

$$X_3 = 4a + 6b$$

If this process continues, total sales will increase over time according to

$$X = (t + 1)\,a + b(t^2 + t) \qquad 0 \leq t \leq \bar{t}$$

This function is obviously convex in t. After the point $\bar{t}$ has been reached the first consumer has exhausted his budget constraint. He settles into his long run steady state spending all his money on the new good. Successive consumers also exhaust their budget constraint. After point $\bar{t}$ has been reached, the aggregate profile is

$$X = (t + 1)\,a + b(\bar{t}^2 + \bar{t}) + (t - \bar{t}) \qquad \bar{t} < t$$

which is linear in t. It is linear because as one new consumer enters the market, another one 'retires' from experimenting and settles into a steady state.

The amount of time taken for a consumer to reach his budget constraint defines his steady state purchases of the good.

$$x = a + b\bar{t}$$

Figure 7.(xiv)

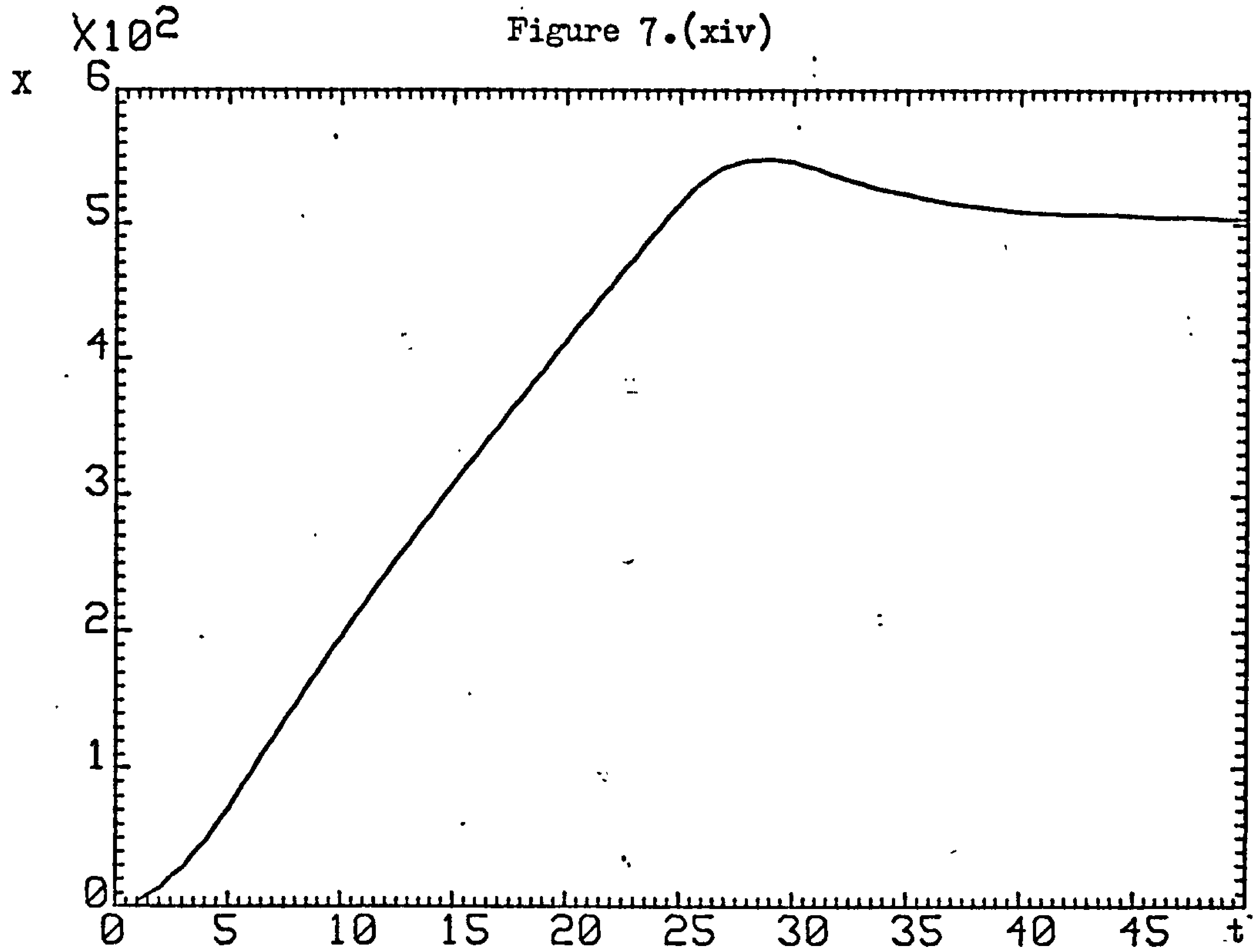
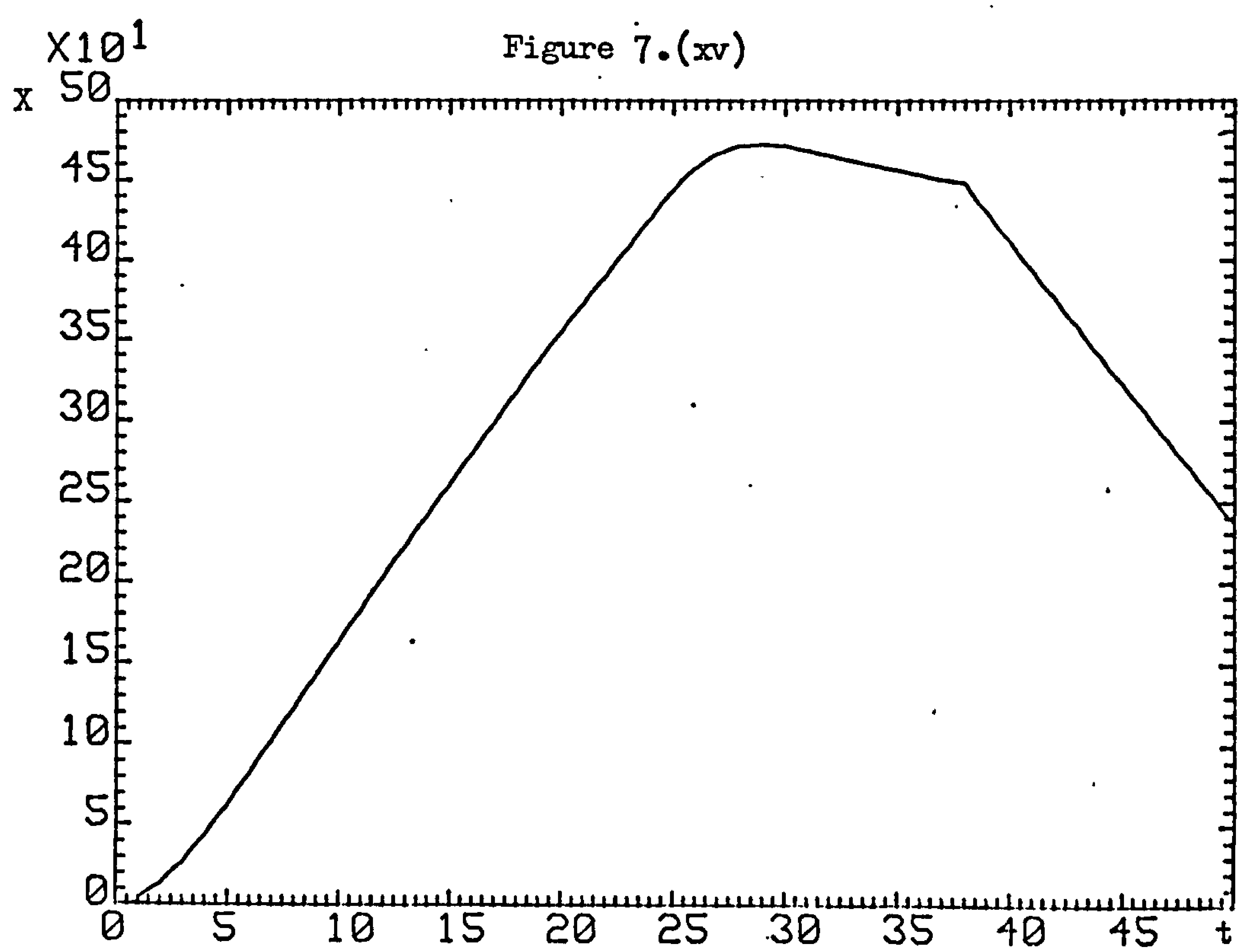


Figure 7.(xv)



Define time t' as the time when the last consumer enters the market, then by time $\bar{t} + t'$, all consumers and hence the market will be in ~~a steady state.~~ Then the steady state market purchases will be

$$X = (t' + 1)\, a + b\bar{t}\, (t' + 1)$$

After point t' has been reached market purchases still increase as existing participants move towards their equilibrium, but at a decreasing rate. The equation for the aggregate time profile becomes

$$X = a(t' + 1) + b\left\{ \bar{t}\,(t' = 1) - \frac{(t' + \bar{t} - t)^2 + (t' + \bar{t} = t)}{2} \right\}$$

$$t > t'$$

It can be seen that the aggregate purchases are concave in t.

These equations hold for a linear individual time trend, but we would also expect them to display the same shape for any steeply concave functions.

We now return to the consumer's decision problem in equation (7.19) and suppose that at each point in time an additional consumer enters the market and all consumers are faced with this decision problem. figs. xiv and xv are the aggregate versions of i and ix, where the number of consumers who enter the market, t', = 25. Both diagrams exhibit the desired shape. They are convex in the early time periods and concave later.

Conclusions

The purpose of this chapter has been to examine the pattern of actual purchases by the consumer over time. We have looked at both the individual and aggregate profiles, and have shown that the aggregate patterns exhibit the desired sigmoid shapes in the early periods; both when the good is adopted, and when it is not. Thus we can offer an alternative rationale as to the shape of the diffusion curve. Instead of explaining it in terms of communication of information between individuals we have argued that individual consumer behaviour alone is sufficient to generate its shape. The knowledge that he is able to update his subjective beliefs about an unknown parameter, affects the consumer's consumption plan. His optimal plan at any point in time is also influenced by his actual experience from the previous periods. It is the combination of his ability to learn and the effect of past learning that generates the required time paths.

The results are obviously dependent upon the form of the utility function and the parameter values chosen. However provided that for an individual consumer, optimal purchases increase fairly quickly over time for a short period at least, then at an aggregate level, the curves in figures xiv and xv will be observed.

It has been assumed throughout this chapter that the behaviour of the firm is relatively passive. The producer of the new good makes an initial advertising statement μ, which is received by individual consumers at different points in time. Thereafter the producer does not take any further part in the process. Given this passive role of the supply side the sigmoid curve is generated by the aggregated consumers' desires to learn about the quality of the new good.

Notes

1. For example, see Baker (1975), Davies (1979), Midgeley(1977) and Gort and Klepper (1982).

2. Davies (1979) suggests an alternative model of diffusion. In a study of 22 process innovations he finds that the shape of the diffusion curve can be explained by differences in firms' characteristics: size, expectations, attitude to risk and type of innovation. Our model suggested here would fit into Davies' more general statement, where consumer differences are due to the point in time at which they become aware of the new product.

Chapter 8

Value of Information

1. Introduction

In the last chapter we showed how a consumer adjusts his optimal consumption plans over time in the light of new information, knowing that he is able to gain additional information for use in the future. It was shown in Chapter 4 that the use of a Bayesian learning process implied an optimal sample size, which enabled the consumer to learn about an unknown parameter, in the most efficient way. By efficient, we mean as quickly as possible.

However, this efficient gathering and updating of information assumes that the process is costless. The computational complexities of estimating an optimal consumption pattern, suggests that there will be costs incurred if only in terms of the consumer's time needed to solve them.

Hey (1981) argues that evaluating utility in every state of nature is no trivial task; when the cost computation of the optimal strategy is taken into account "the optimality of the optimal strategy is no longer self evident".[1]

Hey suggests two alternative strategies for a consumer acting in an environment with imperfect information: suboptimal search, where the searcher searches optimally with respect to a simplified abstraction of the problem and reasonable search, where the consumer searches according to some ad hoc process.

In this chapter we ask: Is there a more efficient learning mechanism? In answering this question we will not attempt to 'cost' the computational difficulties but will look at some simpler decision rules and see by how much the consumer loses out by following these less complex methods. Numerical estimates of the costs of search have been carried out by various authors: Stigler (1961), Gastwirth (1976) and Hey (1981). These studies have the feature that they consider the expected cost of search. It may be that the consumer is interested in other moments of the distribution of search costs.

2. Alternative Learning Mechanisms

Consider a different learning process. The consumer updates his beliefs according to Bayesian rules, but does not bother to look ahead in calculating his optimal purchasing decision. The consumer maximises single period utility every period, using a linear utility function.

$$\text{If} \qquad \mu_t > p \rightarrow x_{t+1} = \frac{M}{p}$$

$$\text{If} \qquad \mu_t < p \rightarrow x_{t+1} = 0$$

But if every x_{t+1} is zero, the consumer will never purchase the new good again, and his information set will never change. In this case it will always be henceforth optimal to purchase the old good. To surmount this degenerate problem, it is assumed that the consumer always purchases one unit of the new good, whatever the expected level of product quality. No matter how bad a product the consumer thinks the new good is, he still experiments with one unit every period. So in the case where $p > \mu_t$: $x_{t+1} = 1$ and $y_{t+1} = M - p$.

The consumer observes the actual level of product quality from the goods or good he has bought and updates his beliefs according to equations (4.15) and (4.16). The general solution to this discrete difference equation is

$$\mu_t = \frac{\mu_0 \phi + \sum_{\tau=1}^{t} x_\tau \bar{\varepsilon}_\tau}{\phi + \sum_{\tau=1}^{t} x_\tau} \qquad (8.1)$$

In the case of a good of which the consumer only purchases one unit per period:

$$\mu_t = \frac{\mu_0 \phi + \sum_{\tau=1}^{t} \varepsilon_\tau}{\phi + t}$$

where ε is the observation on one unit of the good in each period.

. Another type of learning process is where the posterior mean is given by an adaptive process. That is, the difference between the posterior mean and what was actually observed is a linear function of what was expected to be observed and what was observed.

$$\mu_{t+1} - \bar{\varepsilon}_t = \gamma\,(\mu_t - \bar{\varepsilon}_t)$$

$$\therefore\ \mu_{t+1} = \mu_t\,\gamma + (1-\gamma)\,\bar{\varepsilon}_t$$

The solution is

$$\mu_t = \mu_o \gamma^t + (1-\gamma) \sum_{\tau=1}^{t} \bar{\varepsilon}_\tau \gamma^{\,t-\tau} \qquad (8.2)$$

This adaptive function does not depend upon the amount of the good consumed. The weighting given to the initial beliefs relative to experience is determined by an exogenous parameter γ.

3. Comparison of the Learning Processes

We will now compare the value of the three learning processes in terms of opportunity costs of utility foregone. Given a fixed budget constraint, the consumer can only consume one good at the expense of another. Thus the cost of purchasing one good can be considered as the opportunity cost of not consuming the other good. Different learning processes will imply different optimal behaviour patterns, we will now compare these patterns, and see if the consumer is better or worse off using one process instead of another.

The opportunity cost of imperfect information can be measured as we saw in Chapter 3, by the difference between the utility level that could be obtained with perfect information and the utility level that is obtained, due to making decisions with less than full knowledge. These opportunity costs will occur in each period. Over time as the consumer acquires further information these opportunity costs will diminish; but we can compare learning processes in terms of the sum of opportunity costs of utility in each period.

For a linear utility function the consumer will purchase one good or the other depending upon the average quality relative to the price. Let z be the true mean of the underlying normal distribution. Then if $z > p$, the consumer should purchase the new good. If his actual decision is to purchase x_1 of the new good and y_1 of the old good, then his expected utility is $M - px_1 + zx_1$. Thus the expected utility loss EUL is given by:

$$\begin{aligned} EUL &= \frac{zM}{p} - \{ M - px_1 + zx_1 \} \\ &= (z - p)\left(\frac{M}{p} - x_1\right) \\ &= \left(\frac{z - p}{p}\right) y_1 \end{aligned} \tag{8.3}$$

Thus every additional unit of y_1 purchased, the consumer can buy $\frac{1}{p}$th less x_1 and can be expected to lose $z - p$ units of utility. By summing over T periods the expected utility loss L, is given by

$$L = (z - p)\left[\frac{M.T}{p} - \sum_{t=1}^{T} x_t\right] \tag{8.4}$$

Alternatively if $p > z$, then the consumer should spend all his budget each period on the old good, which would yield a utility level of M. By purchasing (x_1, y_1) the consumer can be expected to lose

$$\begin{aligned} EUL &= M - \{M - px_1 + zx_1\} \\ &= (p - z)\, x_1 \end{aligned} \tag{8.5}$$

Summing over T time periods:

$$L = (p - z) \sum_{t=1}^{T} x_t \tag{8.6}$$

These loss functions are estimated numerically. As in the previous chapter we need to make some kind of assumption about the observations actually being generated from a normal distribution with specified mean and variance. It is necessary to generate random numbers since part of the advantage of the Bayesian process over the adaptive is the notion of an optimal sample size, enabling new information to be professed efficiently. Hence we might expect the adaptive process to be more sensitive to random fluctuations, whereas both the Bayesian processes should quickly settle into long run equilibrium.

4. Utility Loss and Risk

The utility loss from any process will depend upon the observations that are taken by the consumer. These observations are stochastic since they are generated by a random variable. Consequently the utility loss from a particular process will itself be a random variable, with an associated probability density function. We can construct the sampling distribution by looking at a series of utility losses from any one process. If the utility loss from process i is L_i, then consumer preferences over the opportunity costs of learning can be represented by a utility function[2]

$$U = f(\tilde{L}_i) \qquad i = 1, 2 \text{ or } 3.$$

Utility is non-increasing in the loss function. The consumer will choose the learning processes with the lowest expected utility loss. If the consumer is risk neutral then he will be concerned with only the mean utility loss, and will adopt the learning process with the lowest mean utility loss. However, if the consumer is risk averse he will be concerned with higher order moments of the distributions. There is nothing inconsistent in supposing that the consumer is risk averse with respect to utility losses from the learning processes, but risk neutral with respect to his choices between the new and old goods (in fact this risk neutral utility function was used to calculate the opportunity costs of each process). We suppose that the consumer has a utility tree;[3] some branches of the tree may be risk neutral, others risk averse. The consumer makes decisions in stages: he first decides which learning process to adopt and then chooses which goods to buy, given his optimal learning process.

We are interested here, in comparing the frequency distributions of the actual utility losses from each process, in order to determine which process the consumer will choose. Tobin (1958) suggests that a consumer's preferences for risky distributions can be summarised by the means and variances. Borch (1969) and Feldstein (1969) show that the means and variances are only sufficient for a statement about risk preferences if the distributions are normal or the utility function is quadratic. Thus if we wish to compare the distributions of utility losses for different processes in terms of only means and variances, the sampling distributions must be normal, or we must be prepared to assume that the utility function is quadratic. Samuelson (1970) argues that the utility function need not be so restricted, provided the frequency distributions are compact. However the sampling distributions generated by the learning processes here, do not seem to satisfy Samuelson's definition of compactness.

A more general approach to preferences between risky distributions is stochastic dominance, developed independently by Hadar and Russell(1969) and Hanoch and Levy (1969). In comparing two distributions, the criteria for first order stochastic dominance is that the cumulative distributions never cross. If the cumulative distribution of utility losses from following process i is $F_i(L)$ and the cumulative distribution from process j is $G_j(L)$, then process i is stochastically dominant over j , for the set of all non-increasing loss functions, if

$$G_j(L) \leq F_i(L) \quad \forall L$$

$$\text{and} \quad G_j(L_k) < F_i(L_k) \quad \text{for some } L_k.$$

The criteria for second order stochastic dominance is that the accumulated area under the cumulative distribution of i is greater than the accumulated area under j . If utility is non-increasing in the loss function at an increasing rate[5], then i dominates j if

$$\int_{-\infty}^{L_k} F_i(L) - G_j(L) \quad dL \geq 0 \ \forall \ L$$

$$\text{and} \quad G_j(L) \neq F_i(L) \quad \text{for some } L_k.$$

Although stochastic dominance is a very powerful and general result, there may be times when the criteria is not satisfied. In which case it is necessary to go back to comparing distributions in terms of preferences for the moments. Scott and Horvath (1980) show that the preference direction depends whether the moment is an odd moment or an even moment. For non-increasing utility functions, the preference direction is negative for positive values of every odd central moment and negative for every even central moment.

We will now apply these criteria to the numerically estimated loss functions.

5. Numerical Results[4]

We quantify the expected utility losses, by evaluating the learning processes numerically. In the terminology of Hey (1981) we might call the learning process examined in Chapter 4, optimal search; the myopic Bayesian learning process outlined in Section 8.2, sub-optimal search; and the adaptive process,reasonable search. The main differences between these three processes is that under optimal search the consumer is aware of his learning capacity; under sub-optimal search, his ability to learn also depends upon the quantity of the new good purchased, though in this case the consumer fails to recognise it. Under reasonable search the ability to learn is exogenous. The computational effort in these three processes, declines from optimal to sub-optimal to reasonable. Parameter values are specified, and then the expected utility loss is calculated for each process over an environment with 80 time periods. Each time this calculation is carried out we will obtain a number for the utility loss from each process. By carrying the calculation a number of times we will be able to build up a sampling frequency distribution of the utility losses from each process. In order to build a sampling distribution which is representative of the population distribution, the utility loss is calculated 500 times for each process. The 500 observations are then divided into classes and a discrete frequency distribution is obtained.

Table 8.1

	μ_o	p	ϕ	γ	M	z	Utility Losses S.O.	R.	O.
(i)	15.5	5	0.01	0.9	200	4.5	59.5	625.0	59.5
(ii)	15.5	5	0.01	0.1	200	4.5	59.5	397.5	59.5

Table 8.1 illustrates the case of an initial high μ_o, but where $z < p$. The result is that both the optimal and sub-optimal processes have the same pattern of purchases, shown in figure 8(i); the consumer initially purchases the new good, but on the basis of his experience immediately reaches the conclusion that $z < p$, and only purchases one unit of the good from then on. However, in the adaptive case, the consumer continues to purchase the new good for a number of periods, making the utility loss relatively high (Figure 8(ii)). A decrease in the parameter γ reduces the utility loss for the adaptive process, as it takes less time for the consumer to adjust to his experience. We can look at the distribution of utility losses from a series of runs of the three processes. Because of the extreme values of the parameters: a high μ_o, which always results in the consumer buying the good initially, and a relatively low z, in the long run the consumer does not buy the good. The two Bayesian processes always result in the same profile of purchases and hence the same utility loss. It is only the adaptive process that exhibits a sampling distribution. This

frequency distribution is shown in Figure 8(iii). The second order stochastic dominance results are given in Table 8.2. The entry in the matrix shows which process is dominant. n.d. means no dominance.

Table 8.2

Process	Sub-optimal	Reasonable	Optimal
Sub-optimal	-		
Reasonable	Sub-optimal	-	
Optimal	n.d.	Optimal	-

Table 8.3 shows the means, variance and degree of skewness of the utility losses for the three processes. Given the second order dominance of optimal and sub-optimal learning over reasonable, and the no dominance result optimal and sub-optimal, due to their distributions being single points, there is little point stating the moments. However this is done to illustrate their irrelevance.

Table 8.3

	Mean	Variance	Skewness
Sub-optimal	59.5	0.0	0.0
Reasonable	416.35	3053.4	33411.4
Optimal	59.5	0.0	0.0

Although the sub-optimal and optimal process yield the same opportunity loss, we might expect that the computational cost of the sub-optimal is cheapest. Hence the sub-optimal process is best.

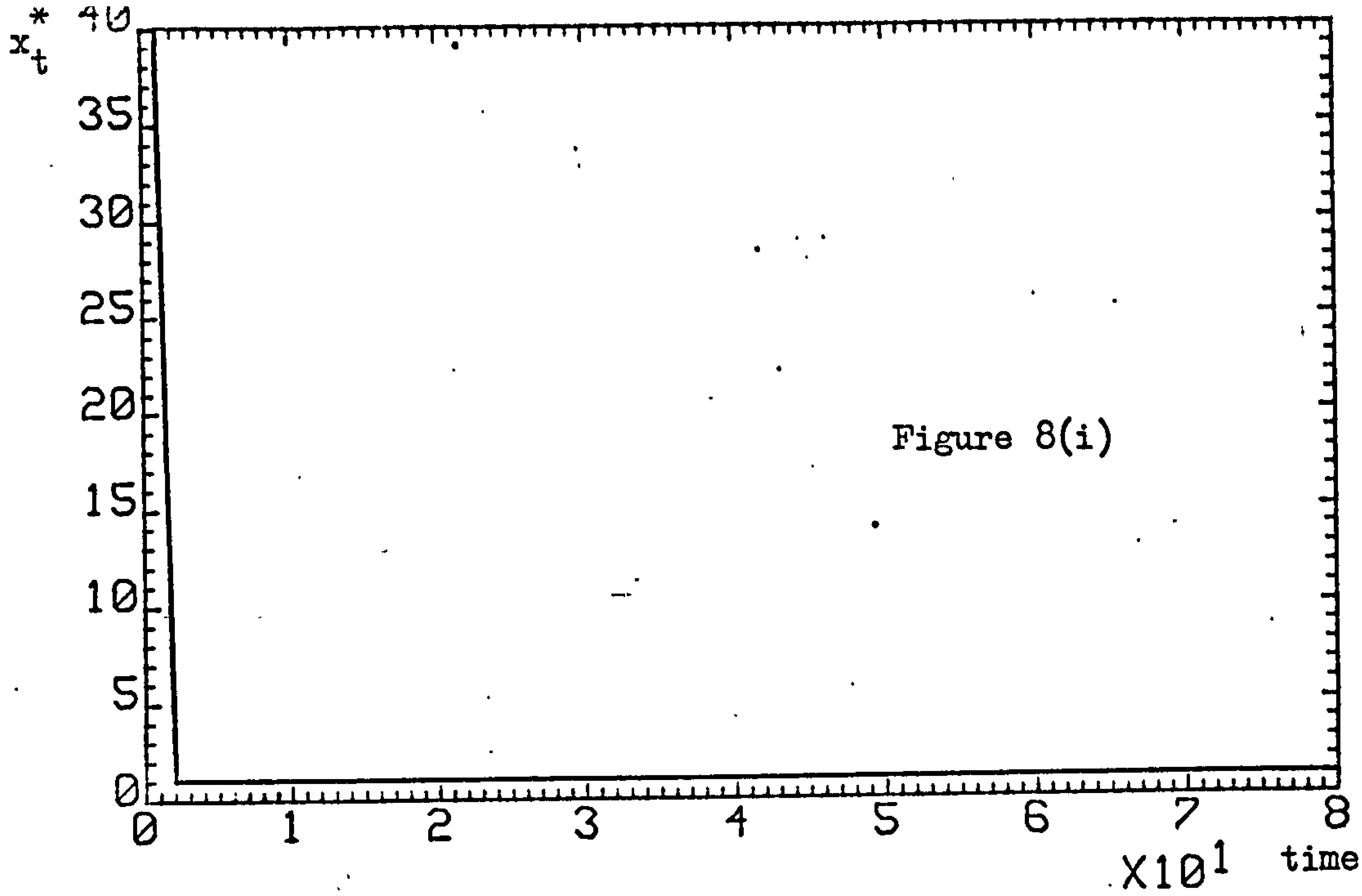


Figure 8(i)

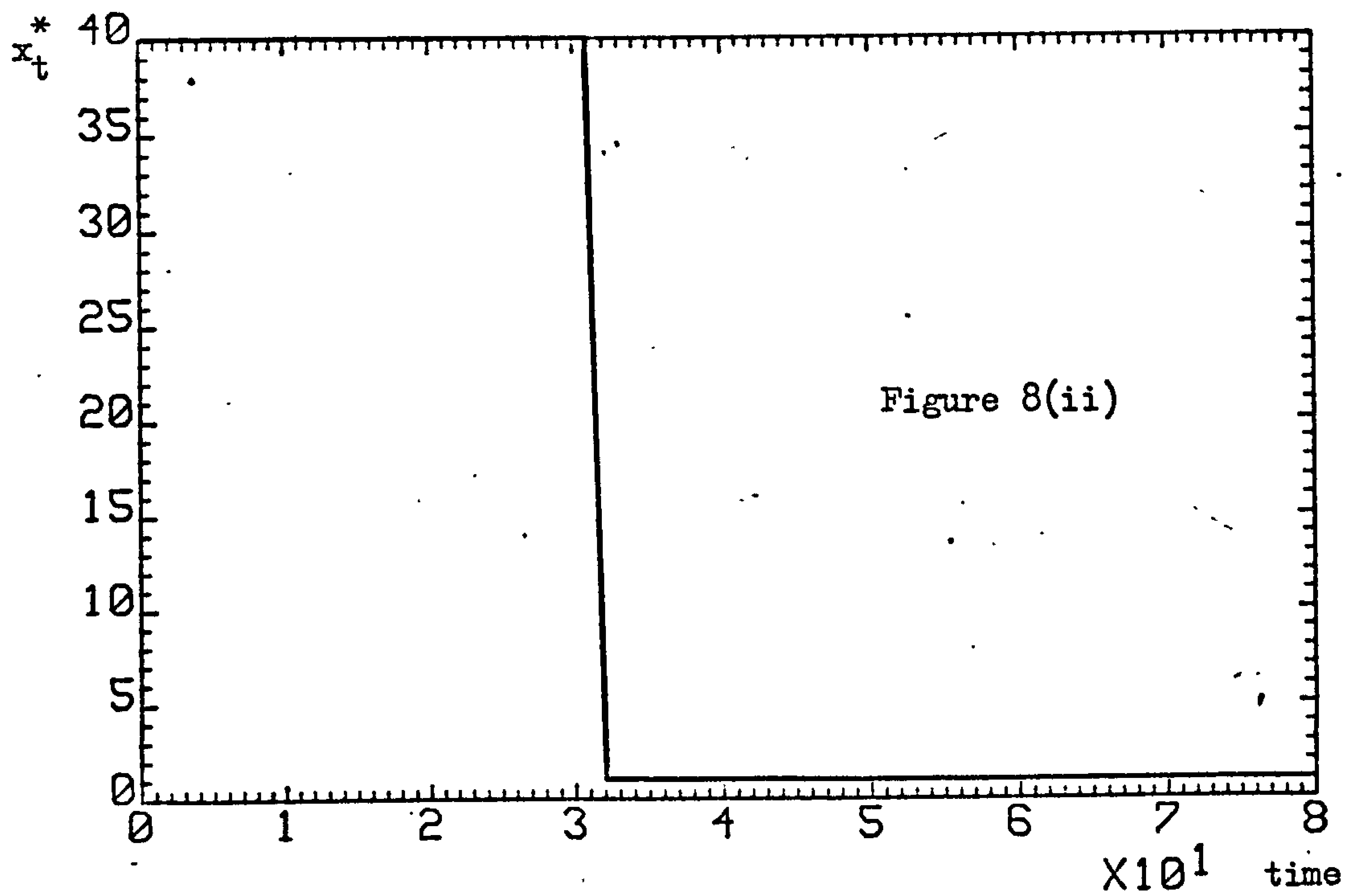


Figure 8(ii)

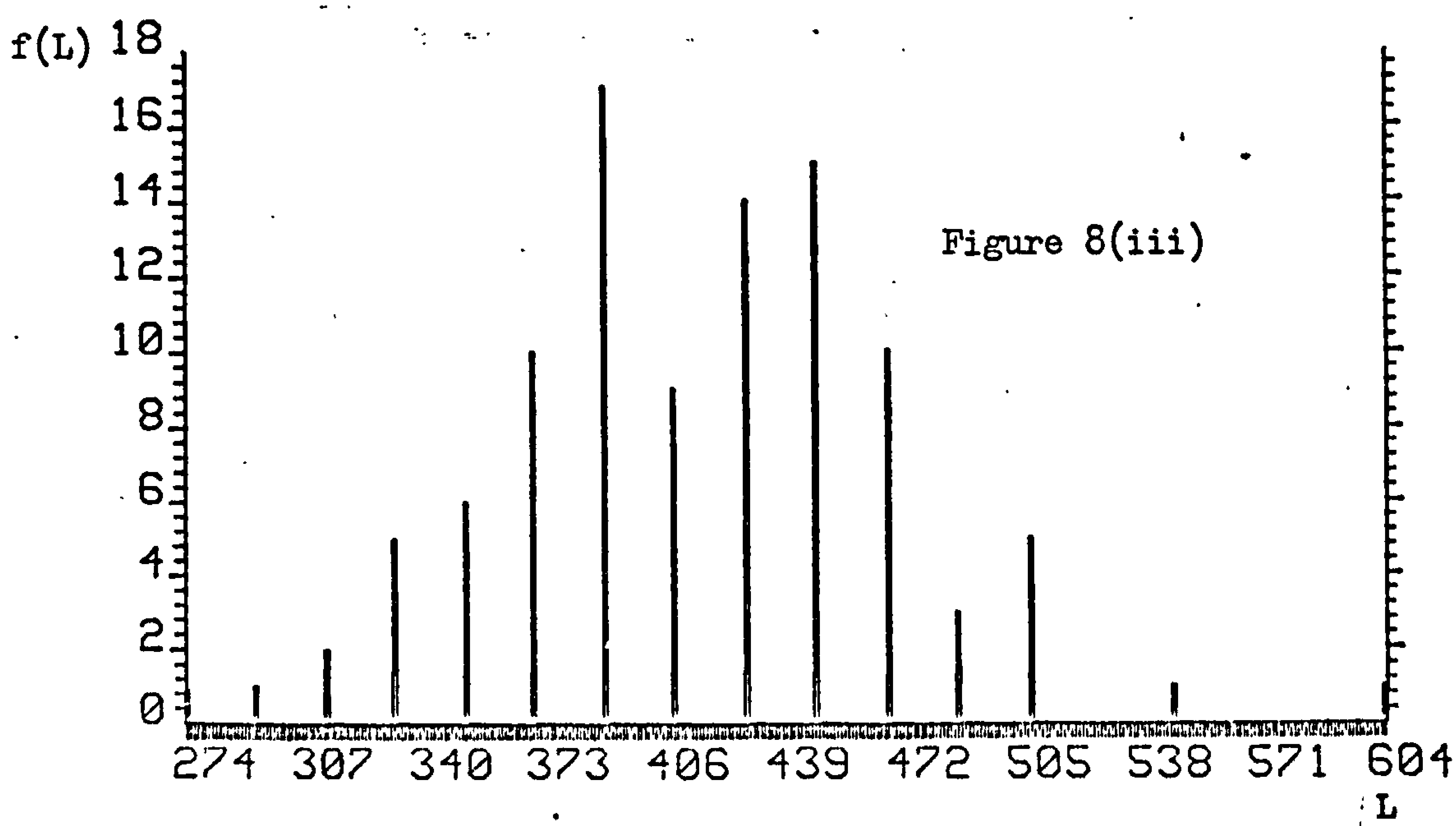


Figure 8(iii)

Table 8.4

	μ_o	p	ϕ	γ	M	z	Utility Loss S.O.	R.	O.
(iii)	15.5	5.0	0.01	0.9	200	4.9	11.9	257.6	15.3
(iv)	15.5	5.0	0.01	0.1	200	4.9	31.4	132.8	16.1

Table 8.4 again compares the difference in utility losses from having different values for γ . z has been increased to 4.9, so that it is now only slightly less than p . The effect of this small difference in $p - z$, is that random samples may occasionally produce the result that the sample mean is greater than p . Figures 8(v) and 8(vi) illustrate the parameters in row (iii). Figures 8(vii), 8(viii) and 8(ix) illustrate row (iv). It can be seen that with $\gamma = 0.9$, the consumer takes a long time to adjust to his experience. But in both cases, as the consumer approaches his steady state expectation he is subject to a great deal of variability due to the closeness of p and z . The two Bayesian processes do not suffer from this variability since they take account of accumulated experience directly, which reduces the impact of sudden observations in the current period. In comparing the two Bayesian processes, the optimal processes in Figures 8(vi) and 8(ix) exhibit internal solutions to the optimal x_t^* in some time periods. This contrasts with the all or nothing optimal values in the sub-optimal cases (Figures 8(iv) and 8(vii), and is due to the ability of the consumer to look ahead in the optimal case, which as shown in Chapter 4, alters the linearity of the utility function.

is (iv) omitted?

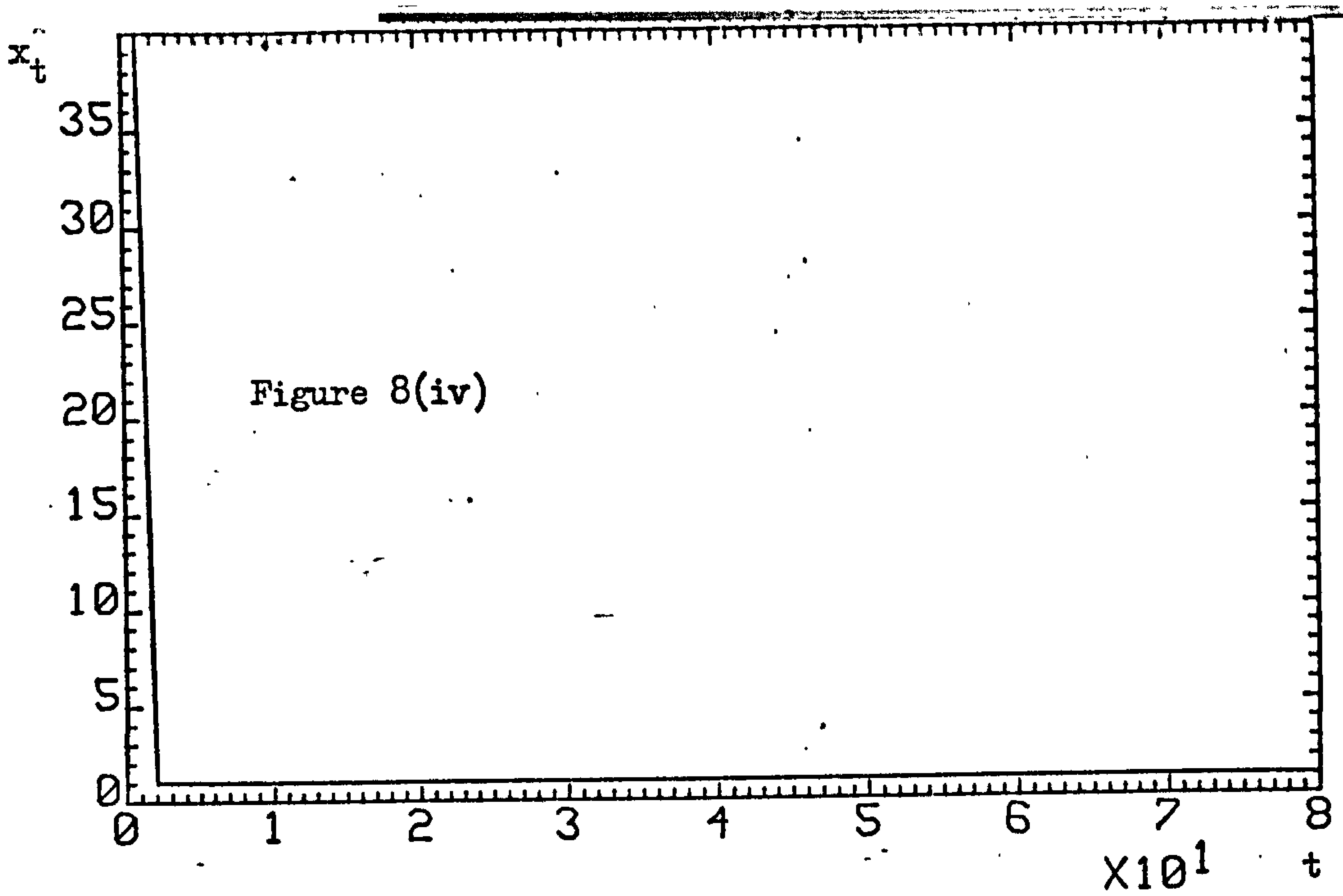


Figure 8(iv)

useful to label diagram

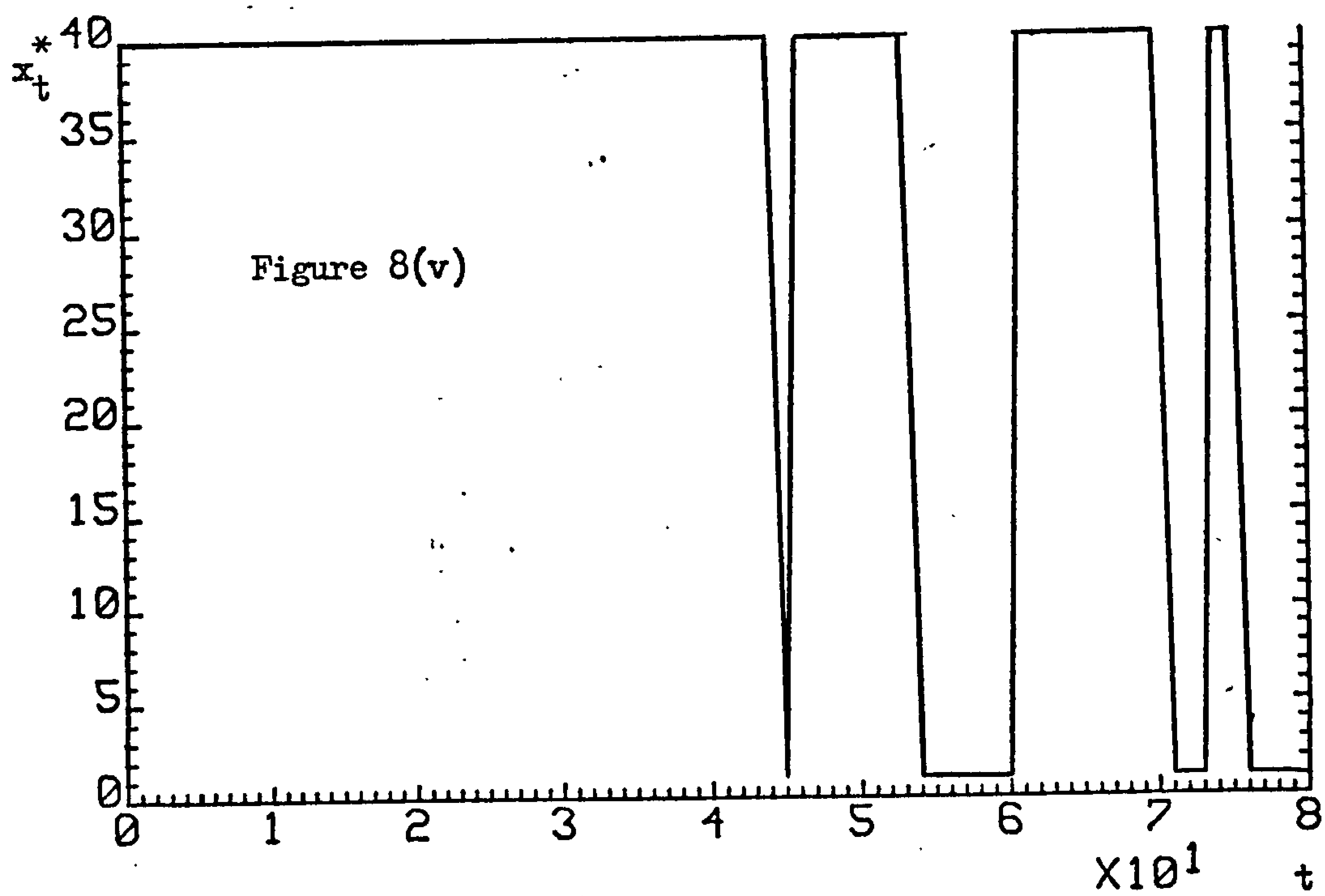


Figure 8(v)

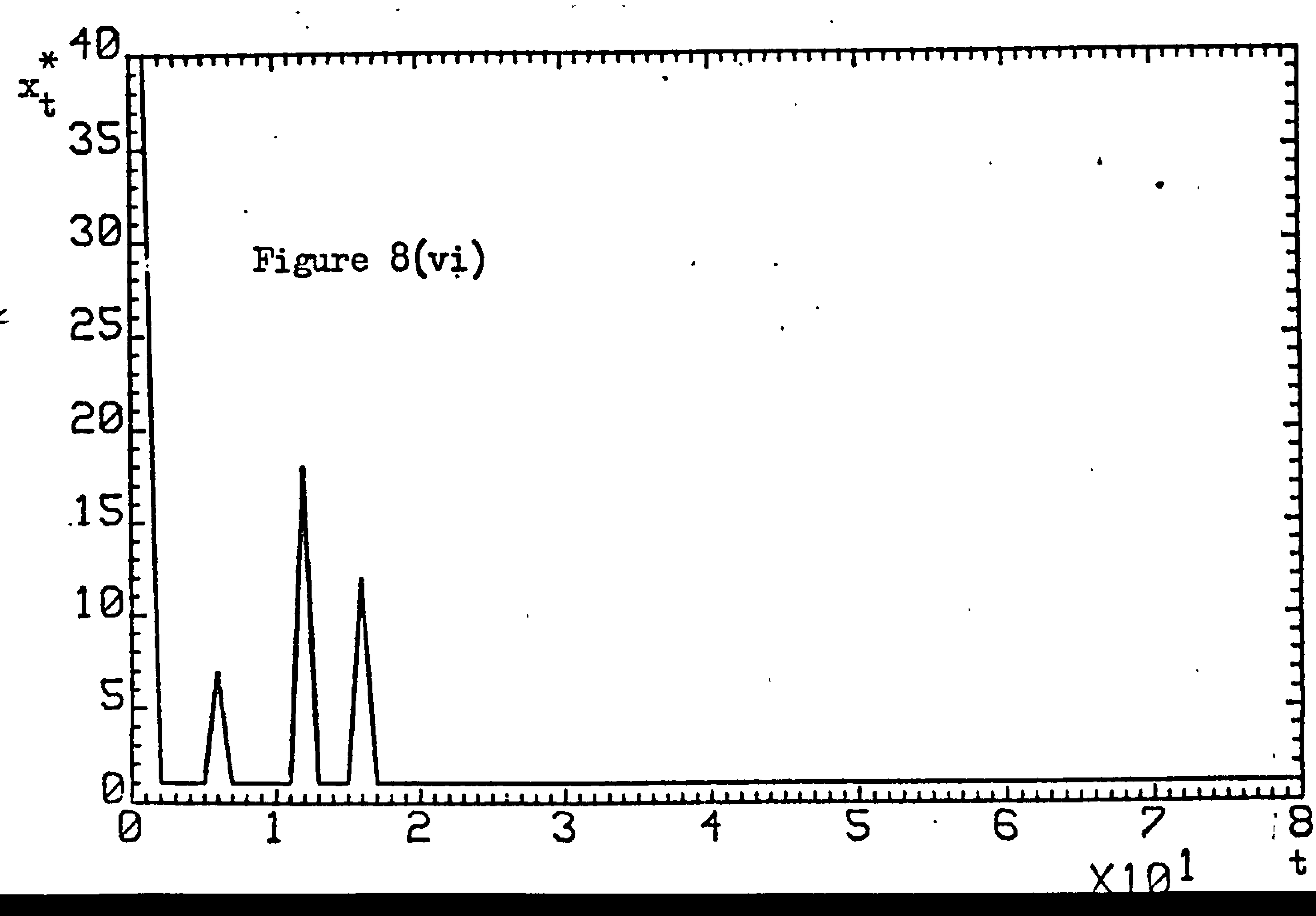


Figure 8(vi)

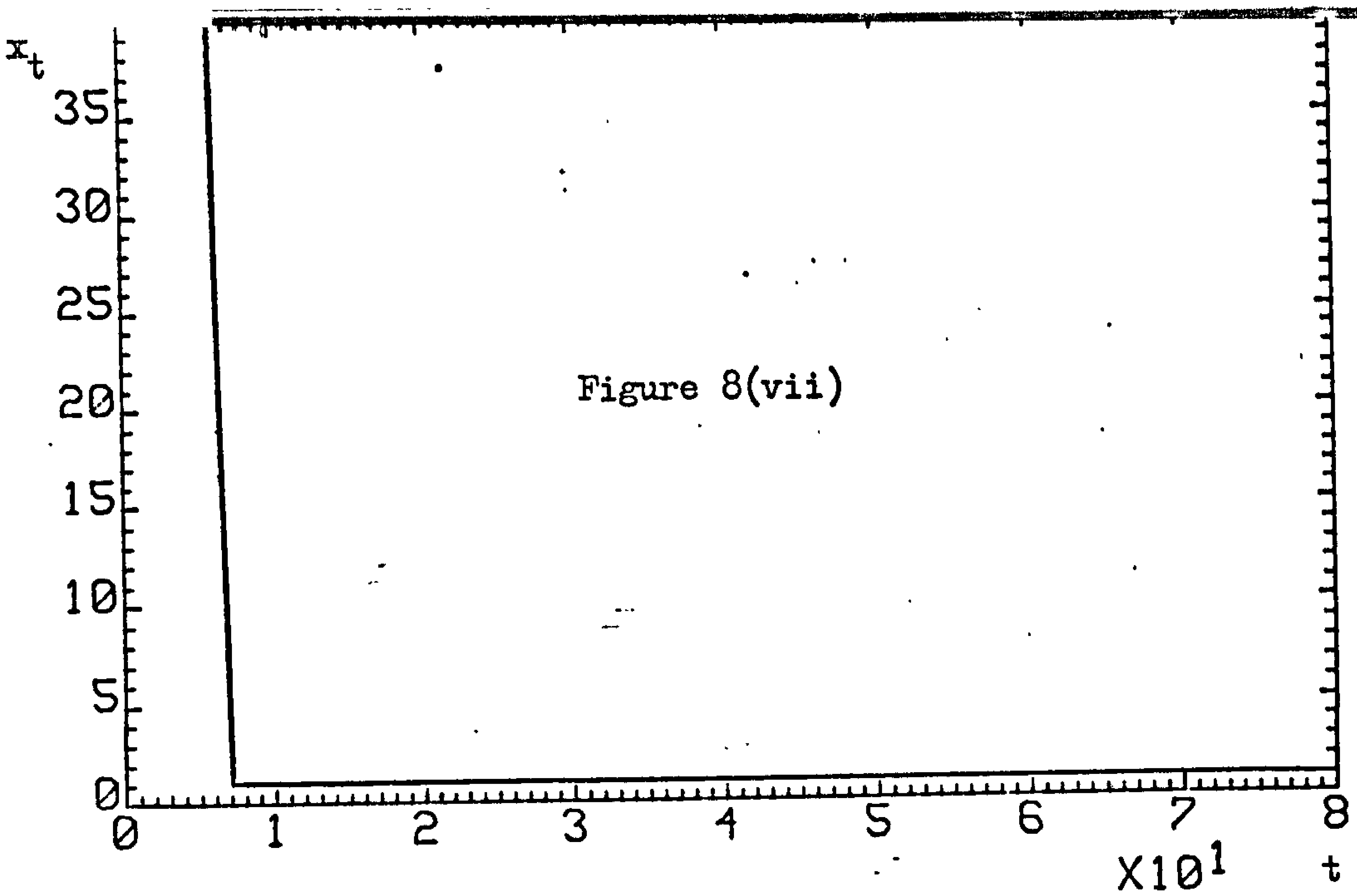
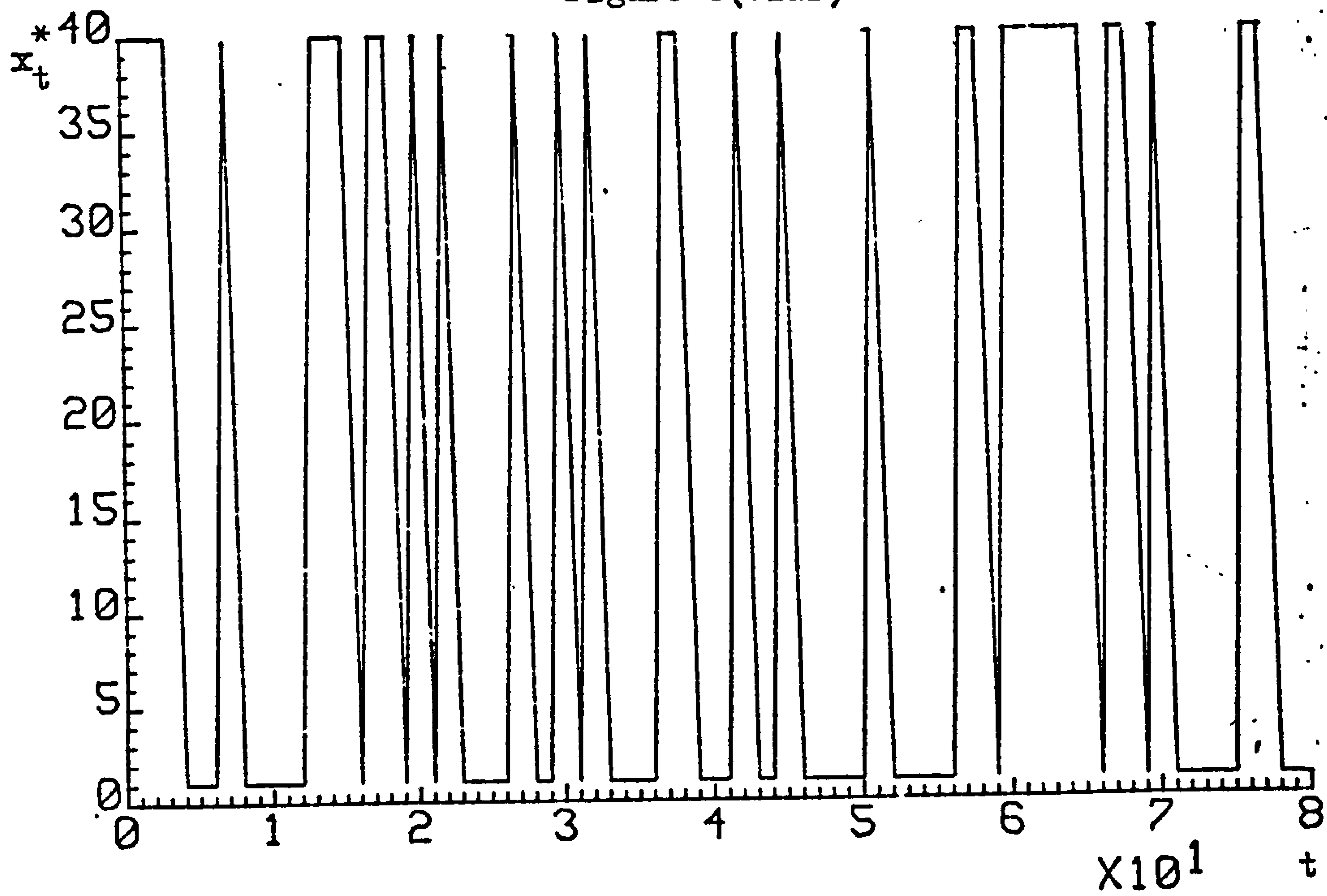


Figure 8(vii)

Figure 8(viii)



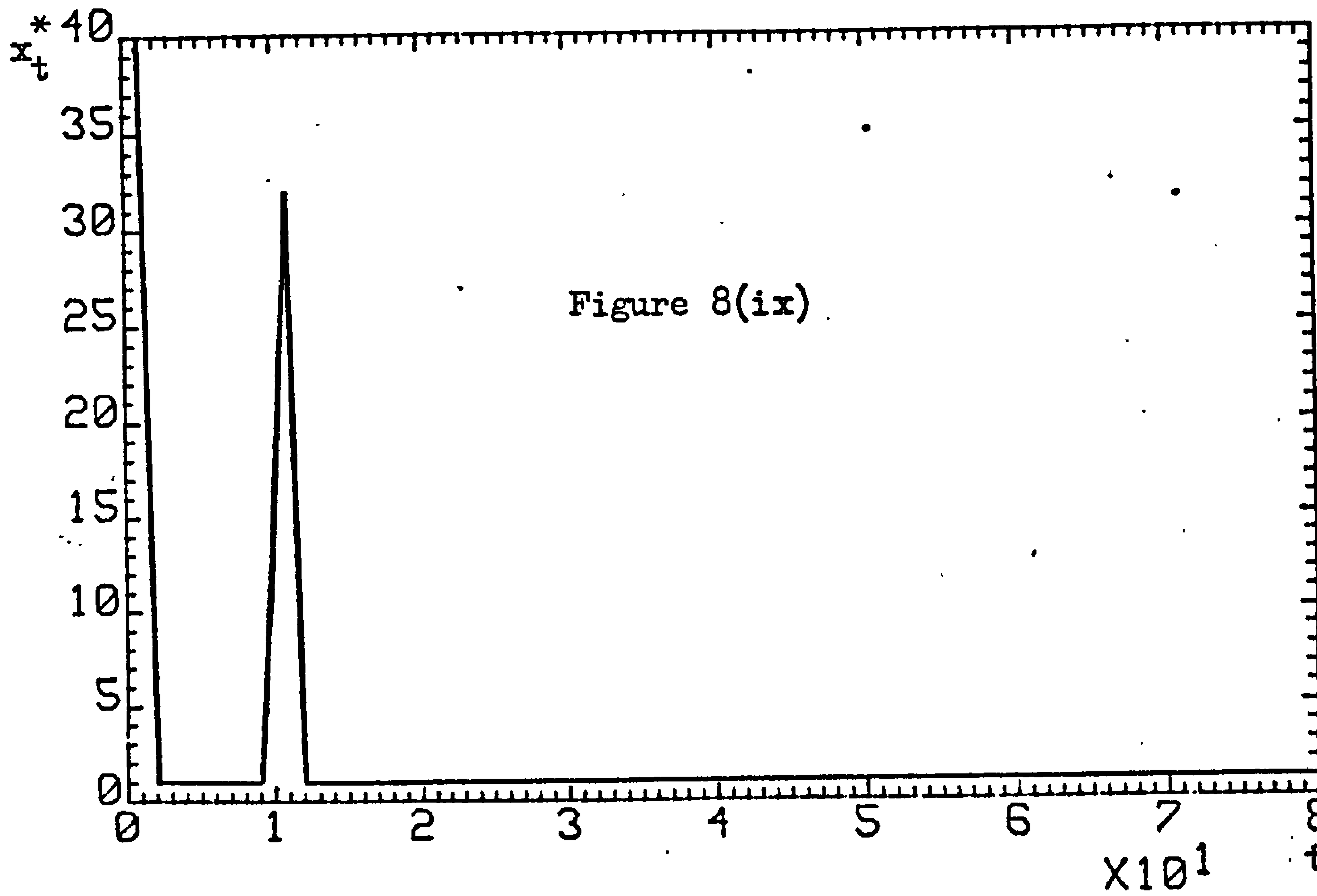


Figure 8(ix)

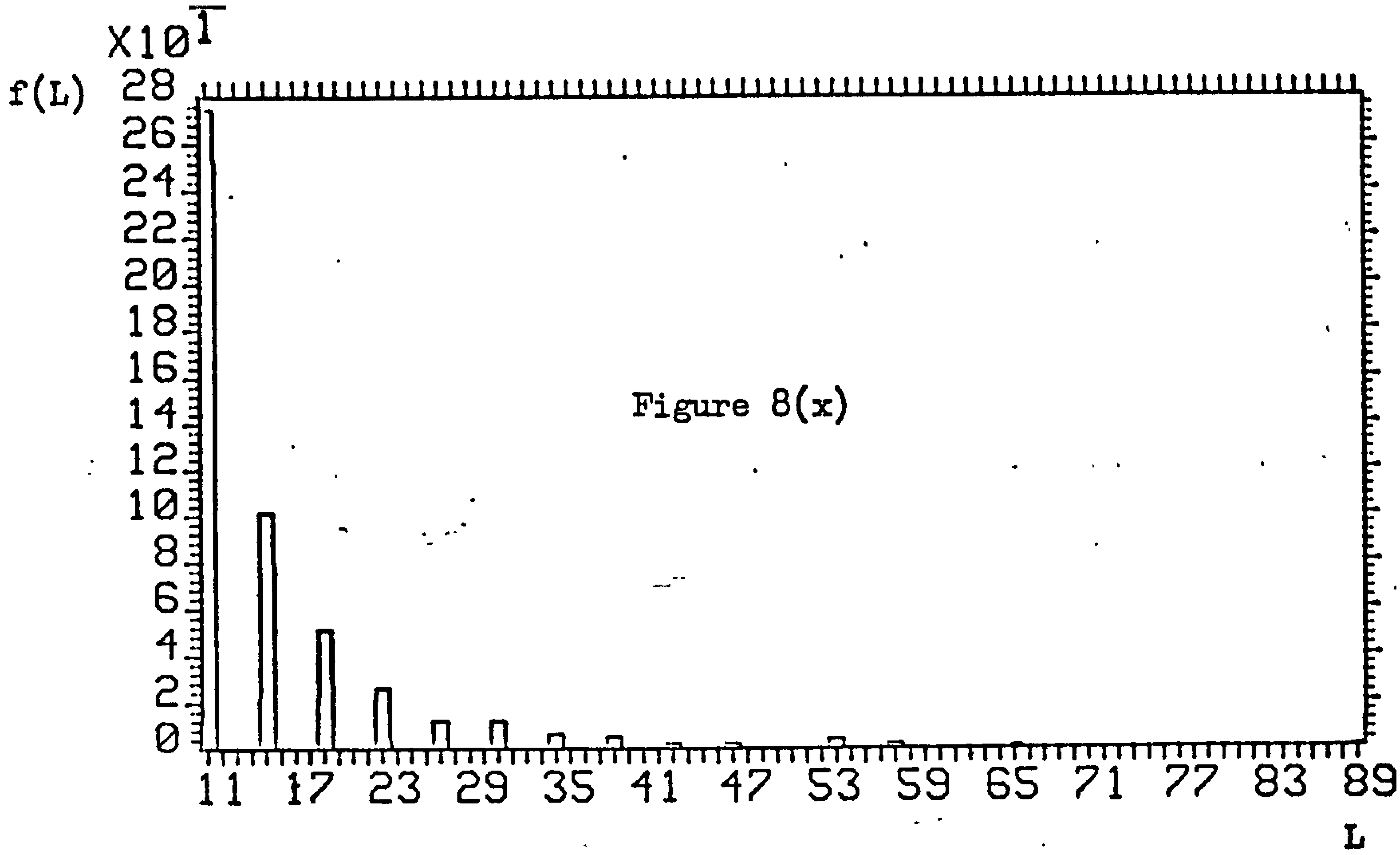


Figure 8(x)

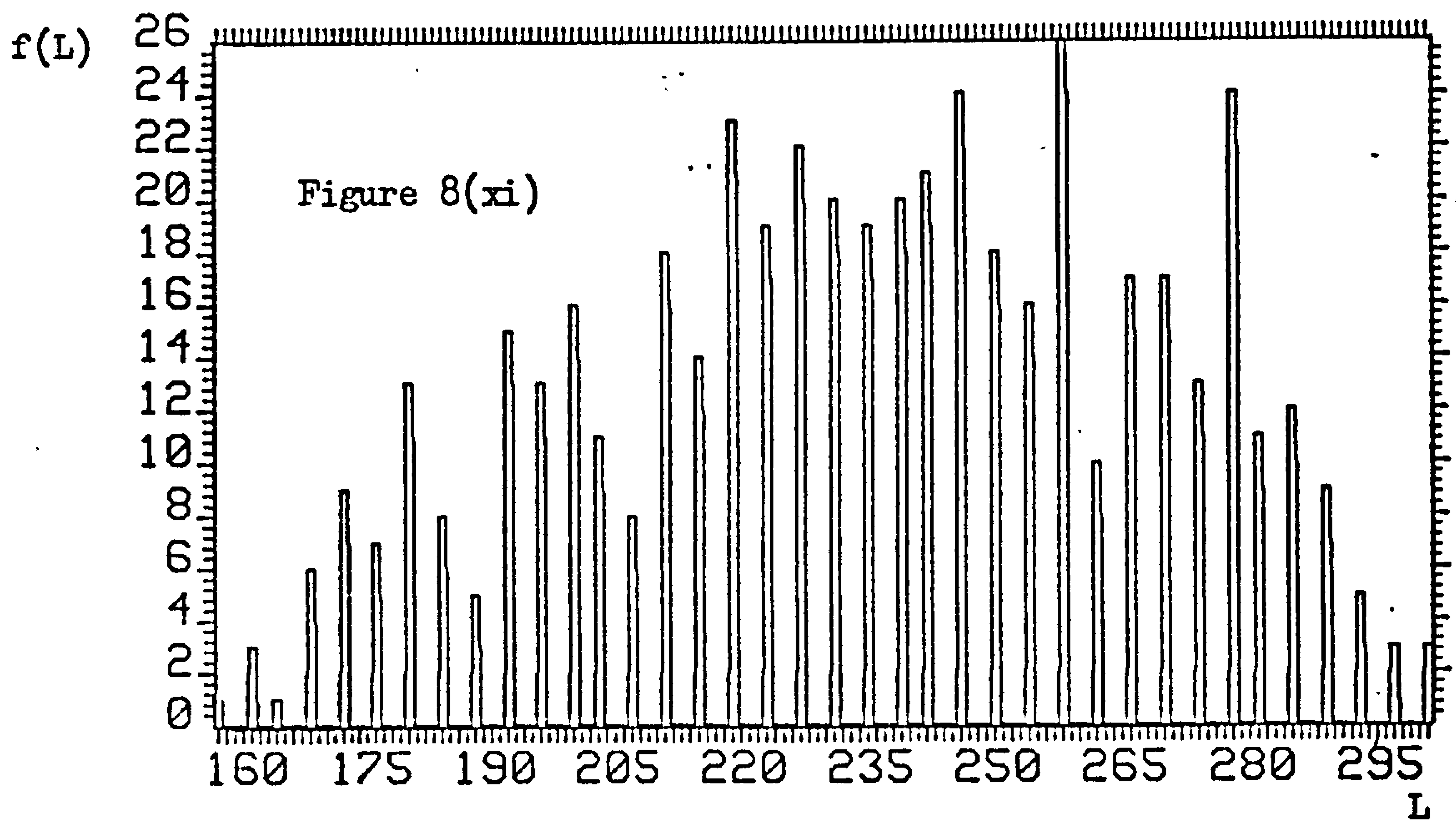


Figure 8(xi)

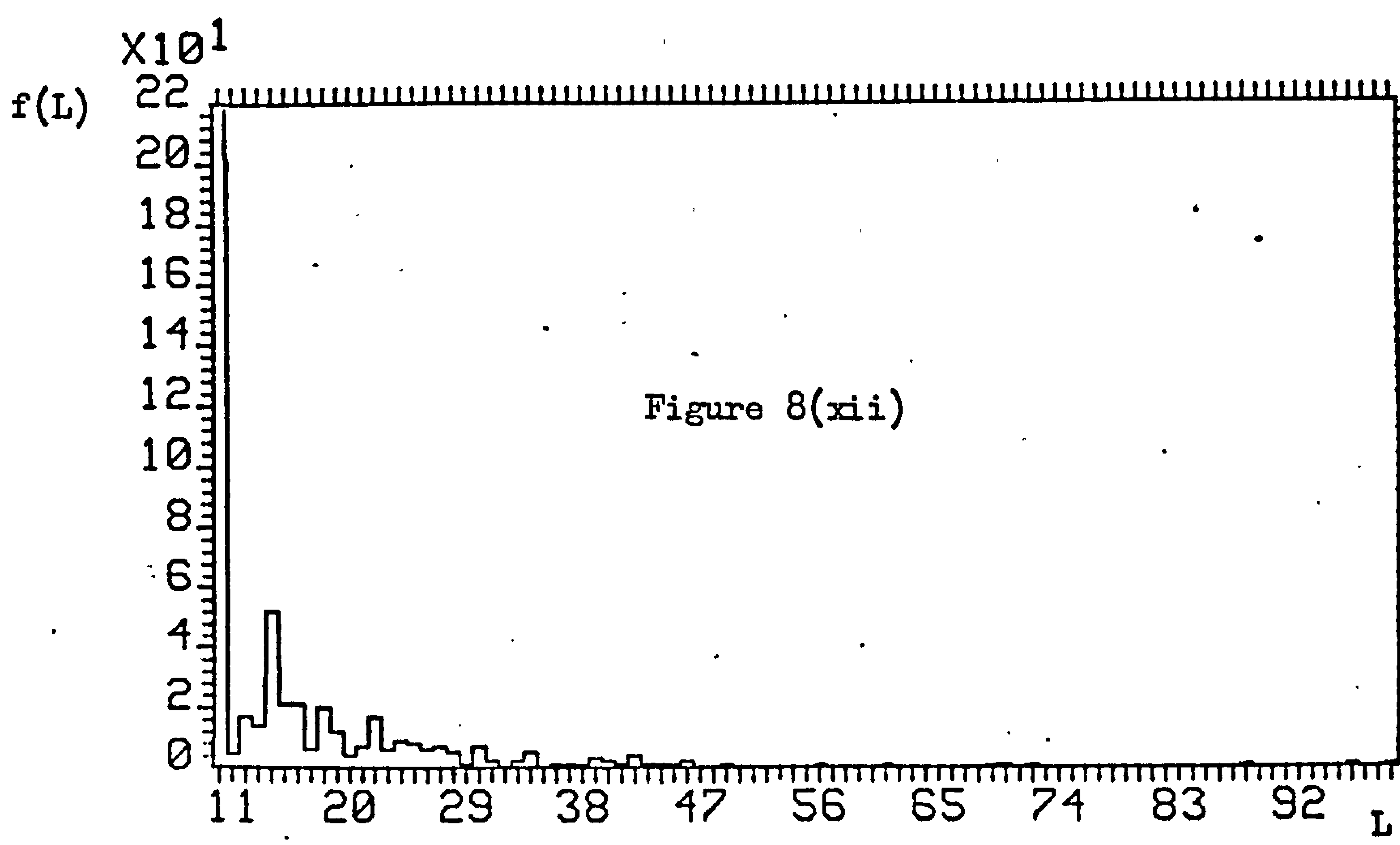


Figure 8(xii)

The frequency distributions for the utility losses for the parameter values in row (iii), are illustrated in Figures 8(x), 8(xi) and 8(xii). The second order dominance results are given in Table 8.5.

Table 8.5

Process	Sub-optimal	Reasonable	Optimal
Sub-optimal	-		
Reasonable	Sub-optimal	-	
Optimal	Sub-optimal	Optimal	-

It can be seen that given the parameters in row (iii) the consumer will be best off following the myopic Bayesian process. The reason that this sub-optimal process is better than the optimal one is that in both cases the consumer quickly learns that product quality is less than price, but in the optimal case, the consumer is tempted to verify this by sampling optimally. Ex post, it can be seen that he would be wiser not to sample.

Table 8.6

μ_0	p	ϕ	γ	M	z	Utility Loss		
4.5	5.0	0.01	0.9	200	4.9	15.8	82.1	48.0

Table 8.6 considers the case where $z < p$ and also $\mu_0 < p$. Thus the consumer starts off with low expectations of quality, and though these expectations increase they are still less than price in the long run. Examples of the profiles of purchases for the three processes are illustrated in Figures 8(xiii), 8(xiv) and

8(xv). These profiles yield the utility losses in Table 8.6 The calculation is computed 500 times to obtain sampling frequency distributions of utility losses for each process and the distributions are shown in Figures 8(xvi) and 8(xviii). The stochastic dominance results are given in Table 8.7.

Table 8.7

Process	Sub-optimal	Reasonable	Optimal
Sub-optimal	-		
Reasonable	Sub-optimal	-	
Optimal	n.d.	Optimal	-

It can be seen that although the reasonable process is inferior to the two Bayesian ones, we are unable to state which Bayesian one is preferred. In order to compare the two Bayesian processes we must have information about the consumer's preferences for the moments of the distribution. The first three moments are given in Table 8.8.

Table 8.8

	Mean	Variance	Skewness
Sub-optimal	15.722	82.37	2293.1
Reasonable	135.577	225.37	327.3
Optimal	15.969	68.65	1418.5

It can be seen that although the sub-optimal process had a lower mean, it has a higher variance and a higher degree of skewness.

It is not always the case that the reasonable process is inferior to the other two. Consider the parameter values in Table 8.9.

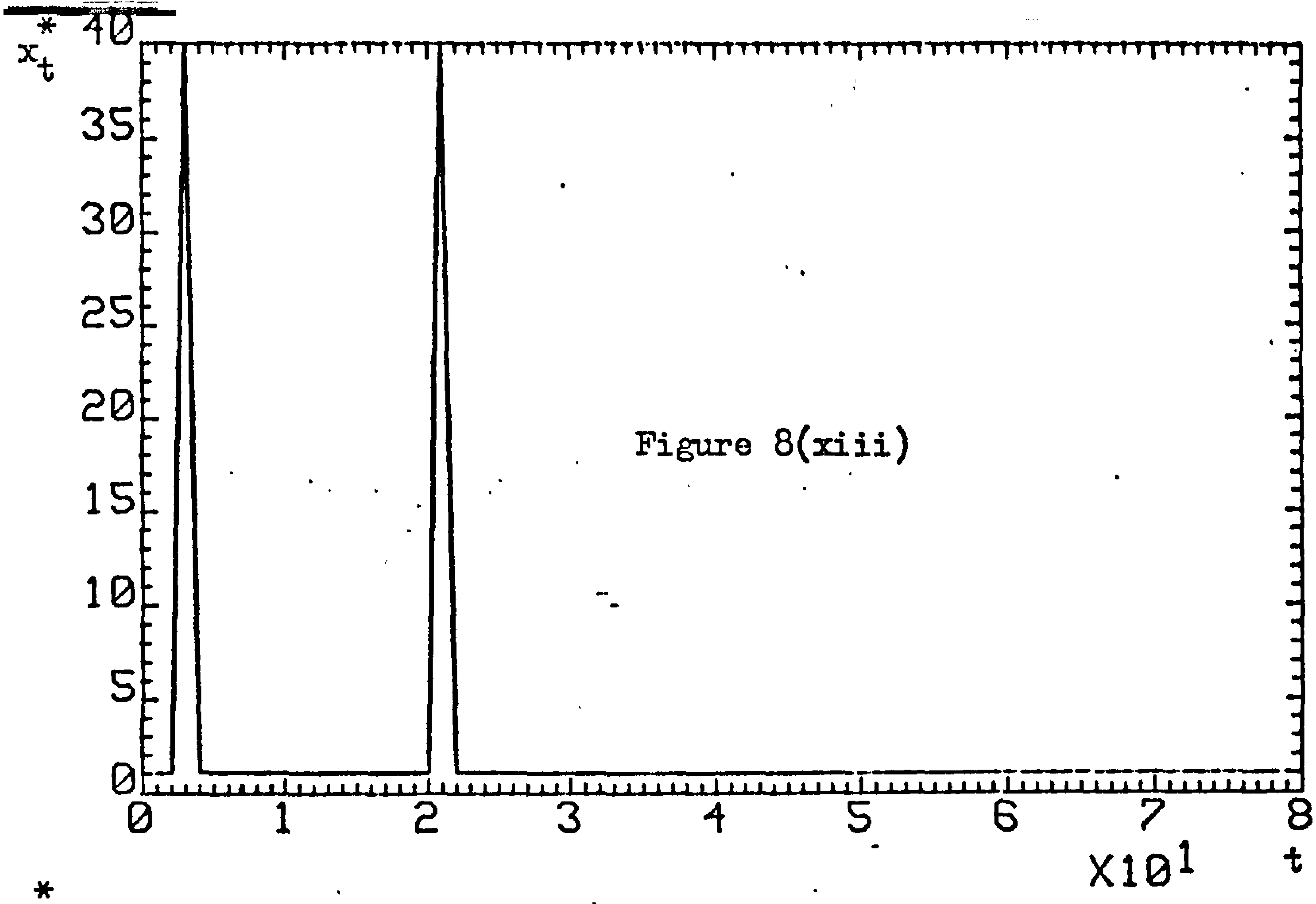


Figure 8(xiii)

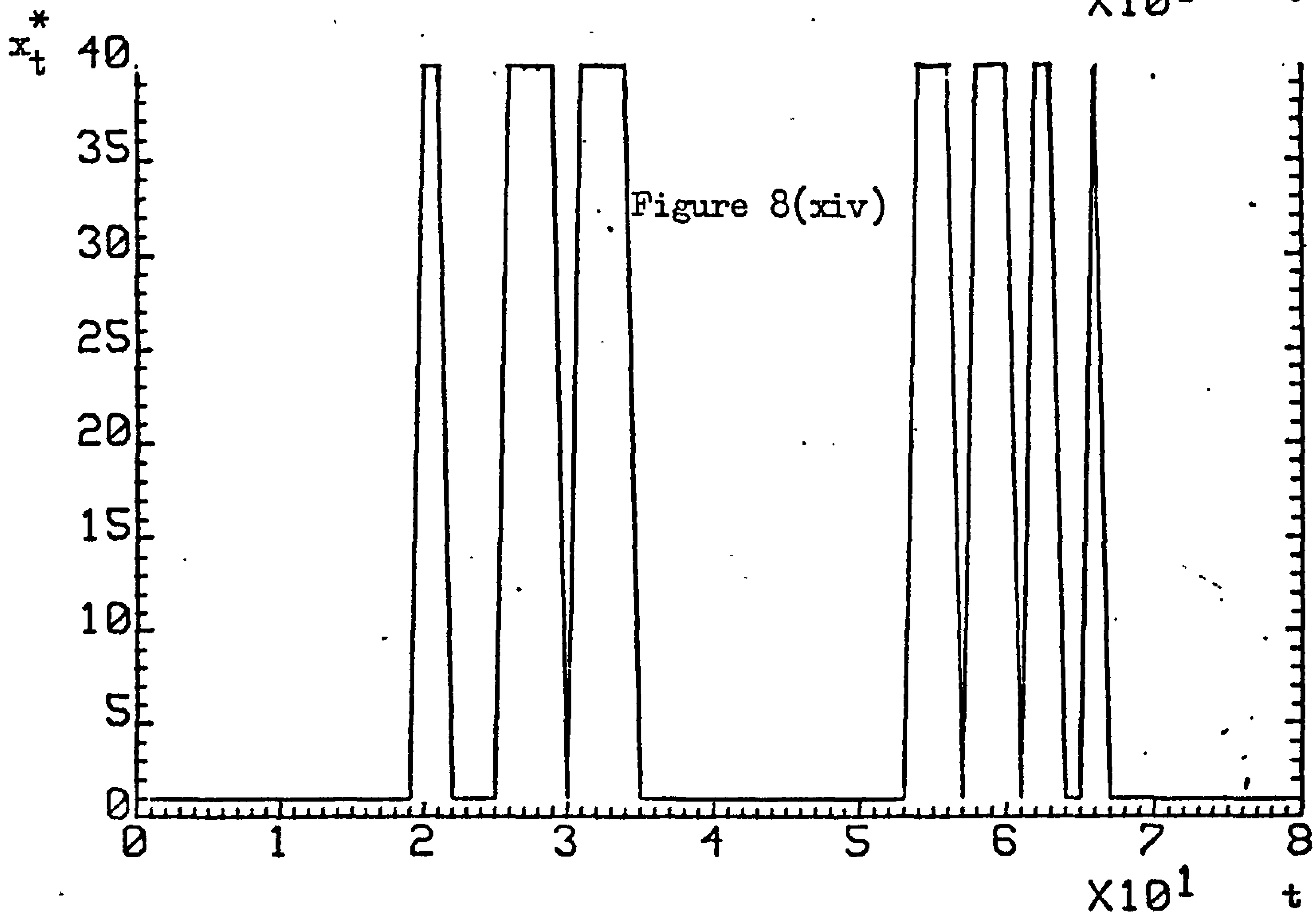


Figure 8(xiv)

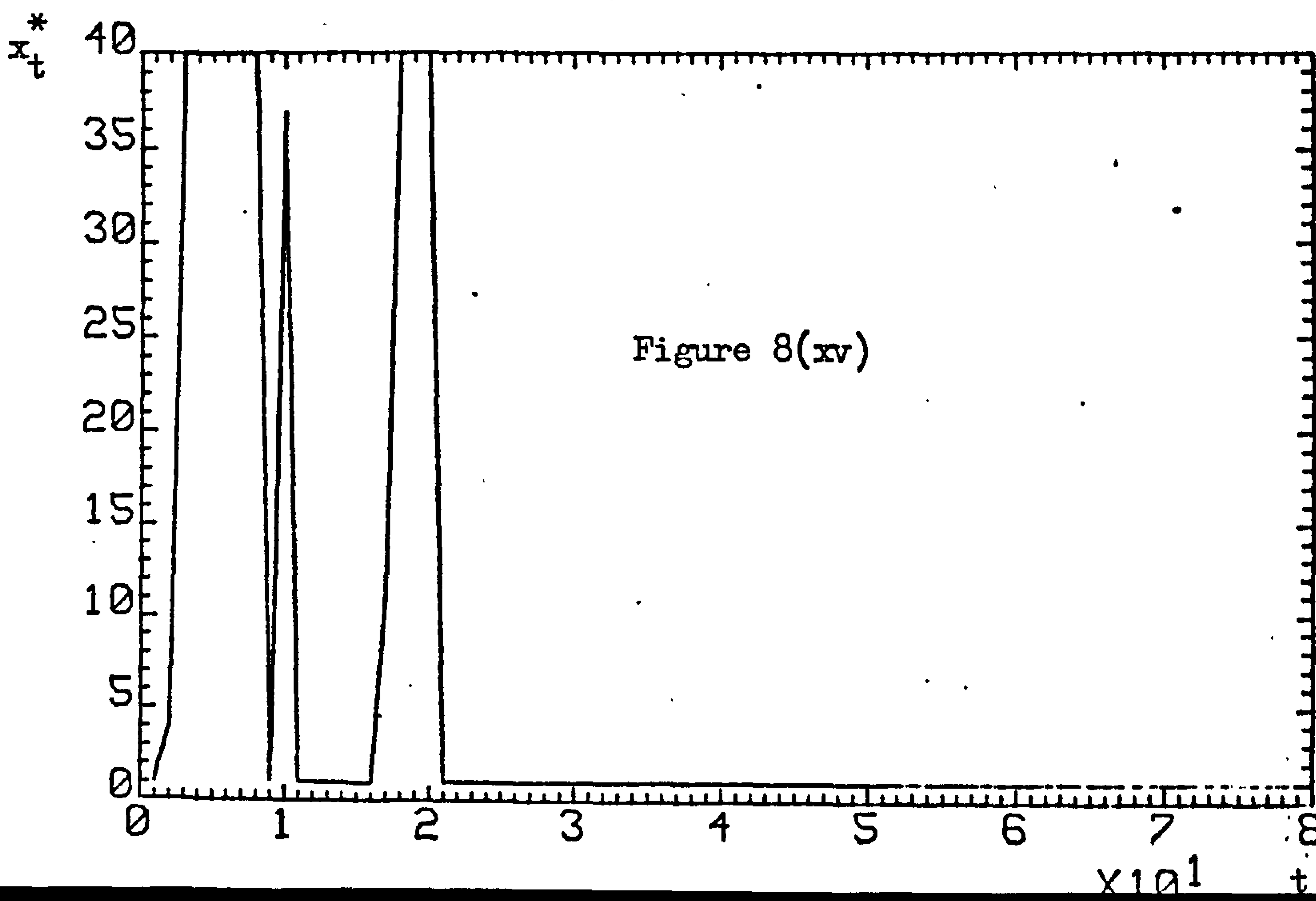


Figure 8(xv)

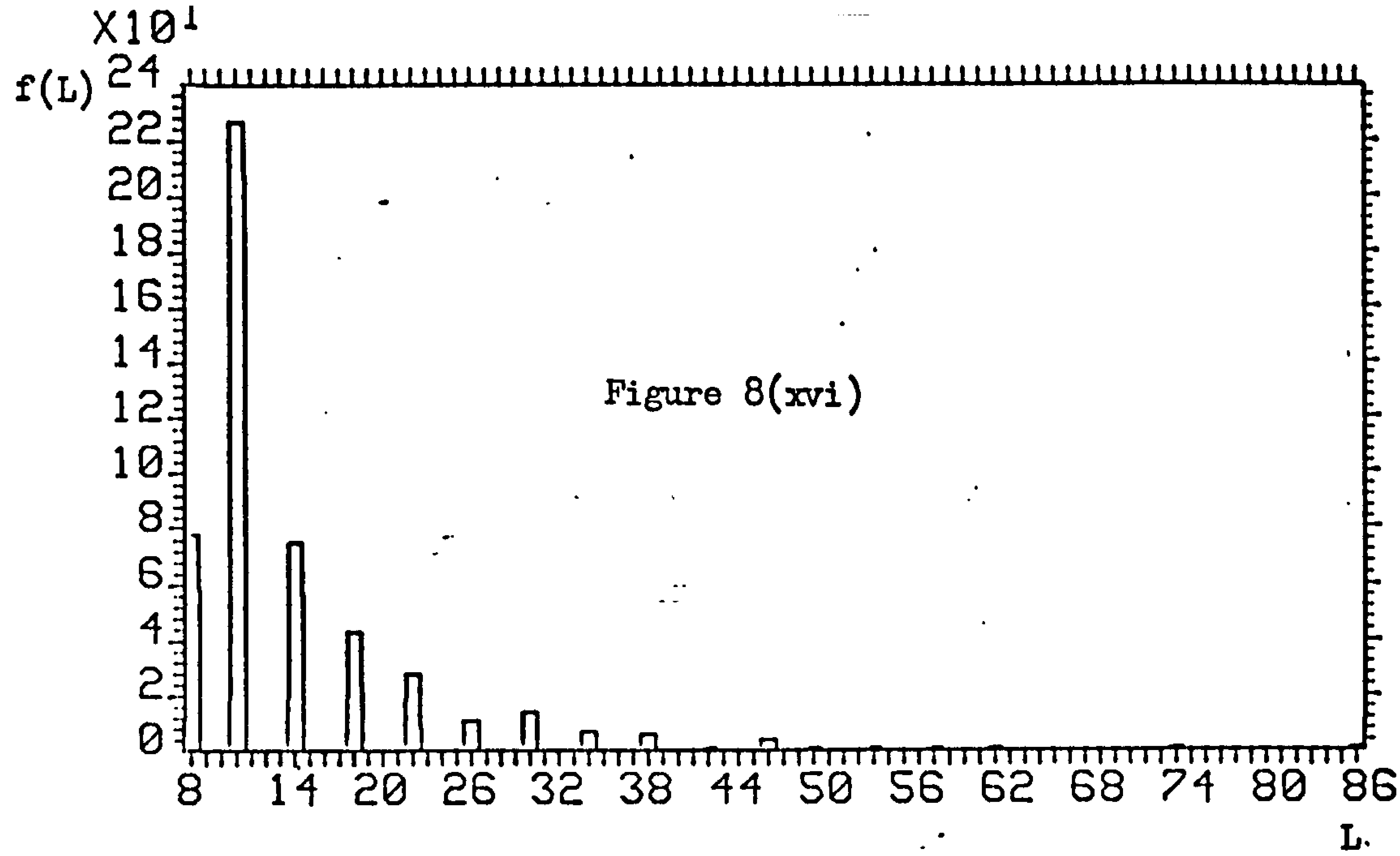


Figure 8(xvi)

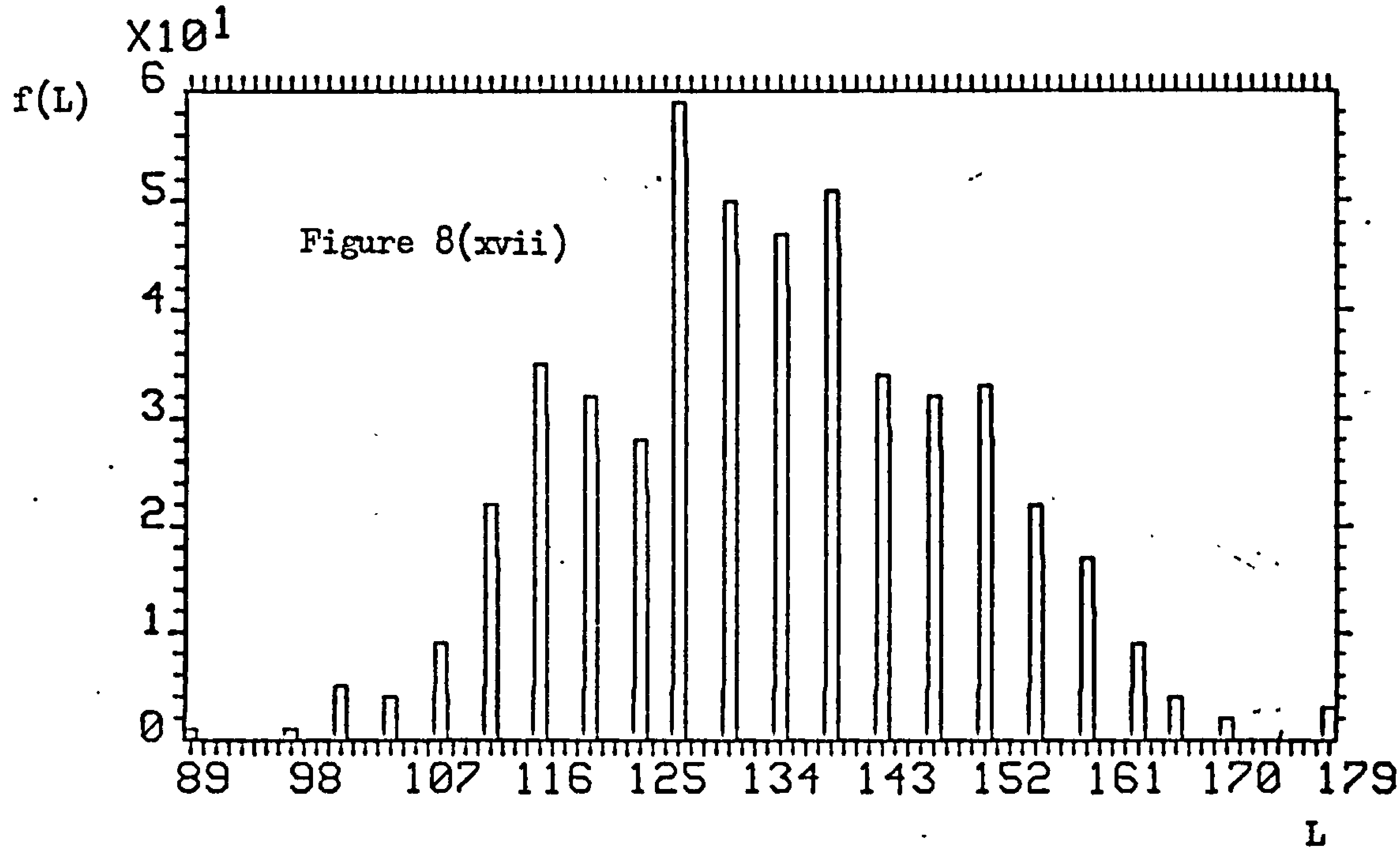


Figure 8(xvii)

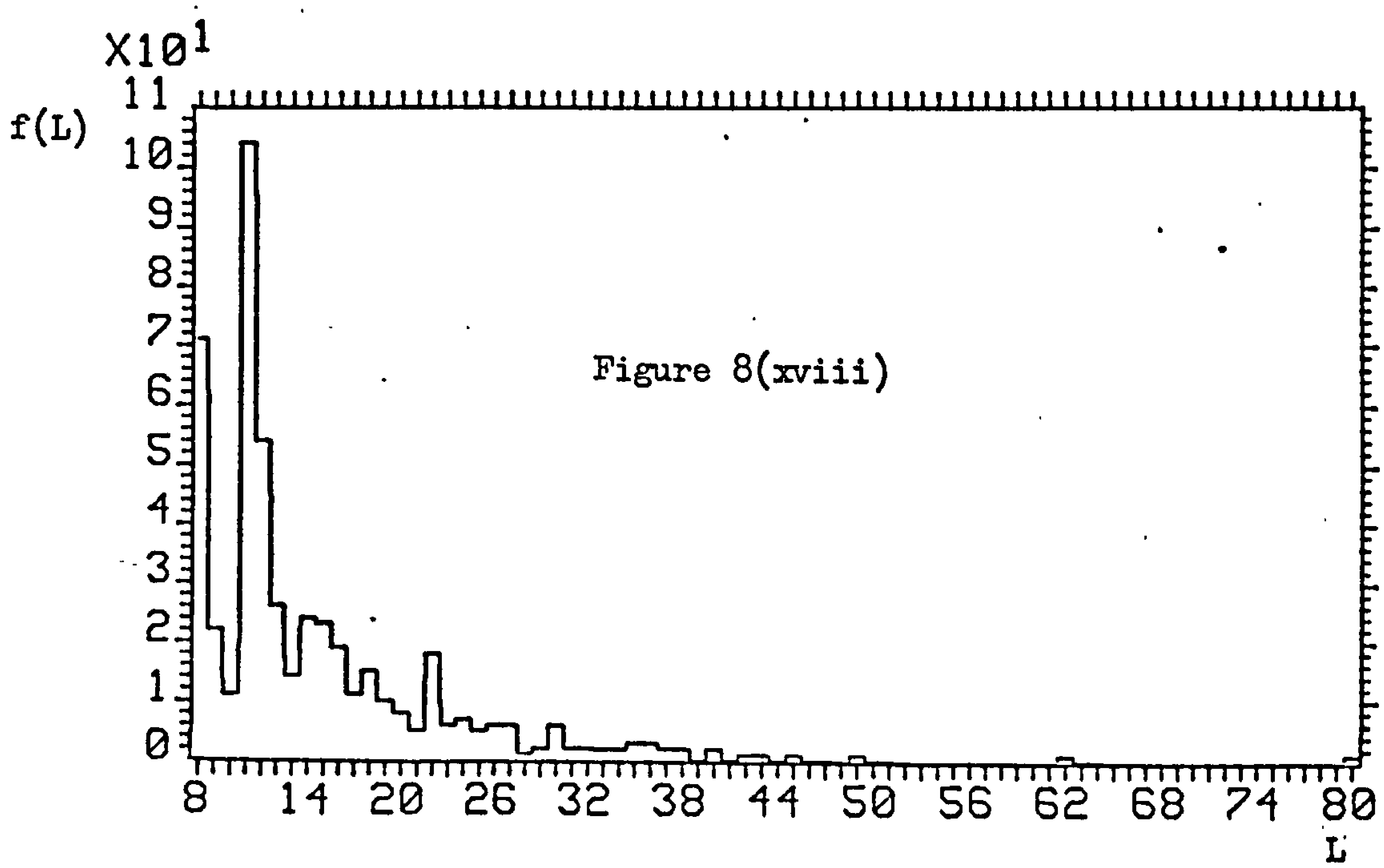


Figure 8(xviii)

Table 8.9

	μ_o	p	ϕ	γ	M	z
(v)	0.2	0.5	0.01	0.9	20	0.4
(vi)	0.2	0.5	0.01	0.9	20	0.6

Row (v) has $z < p$, and row (vi) $z > p$. The stochastic dominance results for each set of parameters is given in Table 8.10 and the values of the first three moments in Table 8.11.

Table 8.10

	z = 0.4			z = 0.6		
	S.O.	R.	O.	S.O.	R.	O.
Sub-optimal	-			-		
Reasonable	Sub-optimal	-		n.d.	-	
Optimal	n.d.	Optimal	-	Optimal	n.d.	-

Table 8.11

	z = 0.4			z = 0.6		
	Mean	Variance	Skewness	Mean	Variance	Skewness
Sub-optimal	16.51	103.7	4879.2	68.89	9895.4	1,437,283.3
Reasonable	120.02	2544.1	3362.1	45.19	1173.4	74,535.8
Optimal	15.61	61.3	1091.2	54.46	8351.1	1,428.116.5

When $z = 0.4$, we obtain the 'normal' result that the reasonable process is stochastically dominated with respect to the second order by the two other processes. Of these two, the optimal process, though not dominant, has a lower mean, variance and degree of skewness. However, when $z = 0.6$, though not stochastically dominant, the reasonable process has the lowest mean, variance and skewness.

The reasonable process also turns out to be the best when $\mu_o > p$ and $z > p$. Consider the parameter values in Table 8.12.

Table 8.12

μ_o	p	ϕ	γ	M	z
1.5	0.5	0.01	0.9	20.0	0.6

The second order stochastic dominance criteria in Table 8.13 shows that the reasonable process is dominant, and the optimal process is dominant over the sub-optimal.

Table 8.13

	Sub-optimal	Reasonable	Optimal
Sub-optimal	-		
Reasonable	Reasonable	-	
Optimal	Optimal	Reasonable	-

Table 8.14 gives the first three moments of the utility loss distributions to illustrate the orders of magnitude of the parameter values in Table 8.12

Table 8.14

	Mean	Variance	Skewness
Sub-optimal	60.41	10,200.7	1,585,627.0
Reasonable	0.36	5.3	105.0
Optimal	42.87	8,444.7	1,604,774.9

The reason for the dominance of the reasonable process is that when both initial expectations and the actual value of product quality are greater than price, under the adaptive learning mechanism the consumer starts buying the new good and continues to do so since the actual quality is greater than price. Because of the large value of γ, the consumer adapts very slowly to new information. The consumer approaches the true value of z from above, and therefore the weighting given to random low observations is offset by the higher initial beliefs. In both the Bayesian processes, the consumer is more likely to be influenced by random low values.

Conclusions

In this chapter we have looked at the costs of following alternative learning processes, in terms of the opportunity costs of utility forgone to the consumer. The three learning processes compared were taken to represent optimal, sub-optimal and reasonable behaviour.[6] All three could be consistent with rationality once an allowance is made for the computational costs of each process. The calculations were evaluated ex poste, that is, we compared the utility losses from each process after the event. We found that no one process always dominated the other two, the best process depended on the parameter values of the underlying utility function, the budget constraint and the true value of the unknown parameter. Of course, ex ante, the consumer does not know the value of the unknown parameter, and is unable to tell which process will turn out to be 'optimal'. The consumer must choose a process to maximise his expected utility, which will involve specifying his attitude to risk.

Notes

1. Hey (1981), p.55.

2. In terms of the value of information discussed in Chapter 3:

$$U = f(U(z,x_z^*) - U(z,x)).$$

We are stating that consumers have a set of preferences over the loss functions.

3. For a discussion of the utility tree and the separability of preferences, see Deaton and Muellbauer (1980).

4. The computer programme used to generate these results is given in the Appendix to this chapter. The stochastic dominance results in this section are all for second order stochastic dominance.

5. It is necessary to assume that the utility function is convex in losses, that is the marginal disutility from a small loss is greater than the marginal disutility of a large loss. This is necessary in order to obtain an unambiguous second order stochastic dominance result.

Notes (continued)

6. In effect each of the learning processes represents optimal behaviour given the computational costs. If costs are zero, then the optimal strategy is indeed "optimal", but for costs greater than zero either of the other two processes may be "optimal", ie the best.

Appendix

```
      PROGRAM AHCPRG
C COMPOSE WITH AHCSUB NETLIB:ROUTIN.PARAM & CSS000:USEFUL.DATAD
      IMPLICIT REAL*8(A-H,O-Z)
      REAL*4 BINS1(700),BINS2(700),BINS3(700),VBEG,VEND
      INTEGER LINE(20)
      LOGICAL*1 OK
      DATA MINB1,MINB2,MINB3/3*9999.9/,MAXB1,MAXB2,MAXB3/3*
     1   -9999.9/
      DATA S1,S2,S3,S4,S5,S6,S7,S8/8*0.0/
      DATA LIMBIN/700/,BINS1,BINS2,BINS3/2100*0.0/,NITER/100/
C---------------------------------------------------------------
C GET RUN PARAMETERS, DISPLAY THEM, AND INITIALISE RANDOM NUMBER
C---------------------------------------------------------------
      CALL PARAM(LINE,OK)
      IF(.NOT.OK)STOP
      WRITE(88,8800)LINE
      READ(88,8810)U0,P,V,D,Z,FM
8800  FORMAT(20A4)
8810  FORMAT(2(F6.3,1X),2(F6.3,1X),F5.1,1X,F6.3)
      WRITE(6,6000)U0,P,V,D,Z,FM
6000  FORMAT('OPARAMTERS (U0,P,V,D,Z,FM): ',4(F6.3,1X),F5.1,
     1   1X,F6.3)
      CALL G05CCF
C---------------------------------------------------------------
C FOR 100 ITERATIONS PRODUCES VALUES FOR "ELOSS" 1,2 & 3
C---------------------------------------------------------------
      WRITE(6,6010)
6010  FORMAT('OITERATION:  ELOSS 1   ELOSS 2   ELOSS 3'/)
C
      DO 500 LOOP=1,NITER
      XP=0.0
      SUM=0.0
      DO 100 J=1,80
      UT=(U0*V+XP)/(V+SUM)
        IF(P.GT.UT)L=1
        IF(P.LT.UT)L=IFIX(SNGL(Z/P))
        XSUM=0.0
        DO 185 M=1,L
          X=G05DDF(FM,0.1D 01)
          XSUM=XSUM+X
185     CONTINUE
        XBAR=XSUM/L
        XP=XBAR*L+XP
        SUM=SUM+L
100     CONTINUE
      IF(P.GT.FM)ELOSS1=(P-FM)*SUM
      IF(P.LT.FM)ELOSS1=(FM-P)*(80.0*Z/P-SUM)
      CALL PUTBIN(ELOSS1,BINS1,LIMBIN,MINB1,MAXB1)
C---------------------------------------------------------------
      ALL=0.0
      XBAR=0.0
      DO 200 IO=1,80
        IF(IO.EQ.1)UT=U0
        IF(IO.GT.1)UT=UT*D+(1-D)*XBAR
        IF(P.GT.UT)K=1
        IF(P.LT.UT)K=IFIX(SNGL(Z/P))
      XSUM=0.0
      DO 285 N=1,K
          X=G05DDF(FM,0.1D 01)
          XSUM=XSUM+X
285     CONTINUE
      XBAR=XSUM/K
      ALL=ALL+K
200   CONTINUE
      IF(P.GT.FM)ELOSS2=(P-FM)*ALL
      IF(P.LT.FM)ELOSS2=(FM-P)*(80.0*Z/P-ALL)
      CALL PUTBIN(ELOSS2,BINS2,LIMBIN,MINB2,MAXB2)
C---------------------------------------------------------------
      SUMIM=0.0
      XP=0.0
      XBAR=0.0
```

```
      DO 300 JO=1,80
        RMAX=-9999.9
      IF(JO.EQ.1)UT=UO
      IF(JO.GT.1)UT=(UT*VT+XSUM)/(VT+IPTR)
      IF(JO.EQ.1)VT=V
      IF(JO.GT.1)VT=VT+IPTR
      IPTR=0
        DO 400 I=1,40
          W=(P-UT)*DSQRT((V+SUMIM)/I)*DSQRT(V+I+SUMIM)
          Y=S15ACF(W,IFAIL)
          RP=0.5*W**2
          IF(RP.GT.150.0)RP=150.0
          A=2.71828**RP
          B=(Z/P)*DSQRT(I/(V+SUMIM))/DSQRT(I+V+SUMIM)
          C=(Z/P)*(UT-P)
          G=Z+(UT-P)*I+Z+C*Y+B*0.3989424/A
          IF(G.GT.RMAX)IPTR=I
          IF(IPTR.EQ.I)RMAX=G
400       CONTINUE
      XSUM=0.0
      DO 385 LO=1,IPTR
        X=G05DDF(FM,0.1D 01)
      XSUM=XSUM+X
385   CONTINUE
      XBAR=XSUM/IPTR
      XP=XBAR*IPTR+XP
          SUMIM=SUMIM+IPTR
300     CONTINUE
      IF(P.GT.FM)ELOSS3=(P-FM)*SUMIM
      IF(P.LT.FM)ELOSS3=(FM-P)*(80.0*Z/P-SUMIM)
      CALL PUTBIN(ELOSS3,BINS3,LIMBIN,MINB3,MAXB3)
C----------------------------------------------------------------
      WRITE(6,6020) LOOP,ELOSS1,ELOSS2,ELOSS3
6020  FORMAT(' ',I7,2X,3F10.2)
      S1=S1+ELOSS1
      S2=S2+ELOSS2
      S3=S3+ELOSS3
      S4=S4+ELOSS1**2
      S5=S5+ELOSS2**2
      S6=S6+ELOSS3**2
      S7=S7+ELOSS1**3
      S8=S8+ELOSS2**3
      S9=S9+ELOSS3**3
500   CONTINUE
C----------------------------------------------------------------
C CALCULATE AND PRINT RESULTS
C----------------------------------------------------------------
      CALL HGRAM(BINS1,LIMBIN,MINB1,MAXB1)
      CALL HGRAM(BINS2,LIMBIN,MINB2,MAXB2)
      CALL HGRAM(BINS3,LIMBIN,MINB3,MAXB3)
      CALL PROBIN(NITER,BINS1,LIMBIN)
      CALL PROBIN(NITER,BINS2,LIMBIN)
      CALL PROBIN(NITER,BINS3,LIMBIN)
C      CALL PRIBIN('BIN1',BINS1,LIMBIN)
C      CALL PRIBIN('BIN2',BINS2,LIMBIN)
C      CALL PRIBIN('BIN3',BINS3,LIMBIN)
      CALL DIFBIN('1-2 ',BINS1,BINS2,LIMBIN)
      CALL DIFBIN('2-3 ',BINS2,BINS3,LIMBIN)
      CALL DIFBIN('3-1 ',BINS3,BINS1,LIMBIN)
      X1=S1/NITER
      X2=S2/NITER
      X3=S3/NITER
      V1=S4/NITER-X1*X1
      V2=S5/NITER-X2*X2
      V3=S6/NITER-X3*X3
      SK1=S7/NITER-3*X1*V1-X1**3
      SK2=S8/NITER-3*X2*V2-X2**3
      SK3=S9/NITER-3*X3*V3-X3**3
      WRITE(6,6040)
      WRITE(6,6050) X1,X2,X3,V1,V2,V3,SK1,SK2,SK3
      CALL EXIT(1)
      STOP
6040  FORMAT('0RESULTS:      ELOSS 1        ELOSS 2        ELOSS 3')
6050  FORMAT('0    X ',3F14.4/'0    V ',3F14.4/'0   SK ',3F14.4)
      END
```

```
      SUBROUTINE PUTBIN(VALUE,BINS,LIMIT,MINBIN,MAXBIN)
C TRUNCATES "VALUE" TO AN INTEGER AND USES IT AS A SUBSCRIPT
C TO INCREMENT ELEMENTS OF "BINS": "MINBIN" AND "MAXBIN" ARE
C THE MINIMUM AND MAXIMUM VALUES OF THAT SUBSCRIPT: "LIMIT" IS
C THE DIMENSION OF "BINS". "BINS" MUST BE REAL*4: "VALUE" IS
C ASSUMED TO BE REAL*8. "MINBIN" AND "MAXBIN" MUST BE
C INITIALISED TO 9999.9 AND -9999.9 RESPECTIVELY, AND "BINS"
C TO ALL ZEROS.
C RAY BURNLEY (SSDPU) 02-JUN-83
      REAL*8 VALUE
      REAL*4 BINS(LIMIT)
      IVAL = IDINT(VALUE)+1
      IF(IVAL.GT.LIMIT) IVAL = LIMIT
      BINS(IVAL) = BINS(IVAL) + 1.0
      IF(IVAL.LT.MINBIN)MINBIN=IVAL
      IF(IVAL.GT.MAXBIN)MAXBIN=IVAL
      RETURN
      END
      SUBROUTINE PROBIN(NOBS,BINS,LIMIT)
C PROCESSES A LOADED "BINS" (DIMENSION "LIMIT") TO GIVE
C CUMULATIVE VALUES OVER CLASSES "MINBIN" TO "MAXBIN" GIVEN
C "NOBS" OBSERVATIONS.
C RAY BURNLEY (SSDPU) 02-JUN-83
      REAL*4 BINS(LIMIT)
      CUMVAL = 0.0
      CNTRIB = 0.0
      OBS = FLOAT(NOBS)
      DO 110 I=1,LIMIT
         IF(BINS(I).NE.0.0) CNTRIB = CNTRIB + (BINS(I) / OBS)
         BINS(I) = CNTRIB + CUMVAL
         CUMVAL = BINS(I)
110   CONTINUE
      RETURN
      END
      SUBROUTINE DIFBIN(LABEL,BINSA,BINSB,LIMIT)
C GENERATES DIFFERENCES BETWEEN ELEMENTS OF "BINSA" AND "BINSB"
C (BOTH DIMENSIONED TO "LIMIT"): "DIF" IS RETURNED EQUAL TO
C ZERO IF THE SIGN OF THE DIFFERENCE CHANGES THROUGH THE BINS
C ARRAYS; OTHERWISE "DIF" IS
C THE AVERAGE DIFFERENCE: +VE SIGN INDICATES "A" > "B".
C RAY BURNLEY (SSDPU) 02-JUN-83
      REAL*8 DIF,WRK,SUM
      REAL*4 BINSA(LIMIT),BINSB(LIMIT)
      IBIN = 0
100   IBIN = IBIN + 1
      IF(IBIN.GT.LIMIT) GOTO 900
      IF(BINSA(IBIN).EQ.0.0.AND.BINSB(IBIN).EQ.0.0) GOTO 100
      DIF = BINSA(IBIN) - BINSB(IBIN)
      SUM = DIF
      NOB = 1
      IF(IBIN.EQ.LIMIT) RETURN
C
      DO 140 I=IBIN,LIMIT
         WRK = BINSA(I) - BINSB(I)
         IF(WRK) 110,140,120
110      IF(DIF) 130,900,200
120      IF(DIF) 200,900,130
130      SUM = SUM + WRK
         NOB = NOB + 1
140   CONTINUE
      DIF = SUM / FLOAT(NOB)
      IF(NOB.EQ.1)DIF=9999999.9
      GOTO 300
C---- SIGN CHANGES SET "DIF" TO ZERO
200   DIF = 0.0
C---- REPORT RESULT
300   WRITE(6,6000) LABEL,DIF
      RETURN
6000  FORMAT('0   DIFFERENCING ( ',A4,' ): AVERAGE...',F10.3,
     &       '            (0.0 INDICATES SIGN CHANGE)')
C---- PROBLEMS
900   WRITE(6,6900) LABEL,IBIN,I
      RETURN
6900  FORMAT('0   DIFFERENCING ( ',A4,' ): ERROR',2I5)
      END
      SUBROUTINE PRIBIN(LABEL,BIN,LIMIT)
      REAL*4 BIN(LIMIT)
      WRITE(6,6000) LABEL,LIMIT
      WRITE(6,6010) (BIN(J),J=1,LIMIT)
      RETURN
6000  FORMAT('0   PRINTING ( ',A4,' ): 1 TO',I4)
6010  FORMAT(/' ',15F8.3)
      END
      SUBROUTINE HGRAM(BINS,LIMIT,MINBIN,MAXBIN)
      REAL*4 BINS(LIMIT),VBEG,VEND
      IF(MINBIN.EQ.MAXBIN)RETURN
      CALL PLOTTE
      CALL DEVPAP(160.0,100.0,0)
      VBEG=FLOAT(MINBIN)-1.0
      VEND=FLOAT(MAXBIN)-1.0
      NCOLS=MAXBIN-MINBIN+1
      CALL HISCHA(BINS(MINBIN),NCOLS,0,1.0,VBEG,VEND)
      CALL DEVEND
      RETURN
      END
***END
?//
```

Chapter 9

Conclusions

In conclusion what can be said about the effect of advertising and learning on consumer behaviour? We have built a neoclassical model of consumer behaviour which allows advertising to be incorporated as an information generating mechanism. The consumer is unsure of the parameter values of a new product and advertising messages provide the consumer with an initial set of beliefs about the unknown elements. The consumer recognises these messages may be biased. Over time the consumer acquires unbiased information by learning about the new good through experiencing it. The importance of the initial advertising message as a component of current beliefs declines over time relative to the impact of experience.

We argued in Chapter 2 that although advertising that affects the consumer's preference ordering, raises difficulties in making welfare comparisons, these difficulties may be overcome by redefining the objectives of preferences. The problem must be redefined so that the new preference ordering is independent of the advertising messages. Advertising affects the consumer only through the implicit budget constraint. In this way, changes in advertising can be compared with changes in prices, and the effects on welfare are easy to compute.

In Chapter 3 we argued that advertising is an information generating mechanism which like all information has an expected value. Whether the consumer seeks further knowledge depends upon the expected value of the advertising message, relative to the cost of the information. Advertising has an actual value in terms of the maximised level of utility obtained from following advertising's advice less the utility obtained in the absence of advertising. We argued that the purpose of advertising is to provide the consumer with prior beliefs about the value of an unknown parameter in the utility function. However, advertising is information supplied by the person selling the product who has an incentive to exaggerate its qualities. The consumer recognises this bias, and in building his set of beliefs, takes account of the inherent deception elements in a series of advertising messages. Having constructed his initial beliefs, the consumer realises that he is able to gain further unbiased information by actually sampling the good.

The implications for behaviour of the consumer's ability to learn about the unknown parameter from his own experience was expanded in Chapters 4, 5 and 6. The differences between these chapters related to the different assumptions made about the environment. In Chapter 4 we considered the case where the consumer had a linear utility function, normally distributed prior beliefs, and the underlying random variable has a normal distribution. The important result from this chapter, repeated in the next two chapters, was that the opportunity for learning, in terms of updating his beliefs about unknown parameters does affect consumer behaviour. If the information enabling him to carry out the updating procedure is a function of the amount of the good consumed, then the consumer will purchase more of the good than if the opportunity for learning did not exist.

This conclusion was shown to be reversed for the case of a two period budget constraint and a strong preference for consumption in the first period. Finally in Chapter 4, we considered the implications of sampling on the efficiency of the household technology, and showed that a single period "inefficient" solution could exist. Of course this solution is optimal over two periods.

In Chapter 5 we considered the case of a constant elasticity of substitution utility function where the underlying random variable has a log normal distribution. This enabled us to see the effect of risk aversion on the problem. The result, not unexpectedly, was that as previously the consumer purchases a greater quantity of the decision variable in the first period, enabling him to learn optimally about his environment.

In Chapter 6 we considered the case of a nonsymmetric prior, taken to represent the consumer's inherent suspicions as to the truth of an advertising statement. We derived the posterior distribution for this prior and a random sample from a uniform distribution. As previously, the consumer purchases a larger quantity of the new good due to sampling.

The increased purchases of the decision variable in the first period in an adaptive environment can be contrasted with the non-adaptive two period models of Sandmo (1970) and Modigliani and Dreze (1972), where, provided certain conditions are satisfied the presence of uncertainty in the second period leads to a lower optimal value of the decision variable in the first period. The risk averse consumer 'saves' for the possibility of a run of bad luck next period. With adaptive distributions these conclusions are reversed.

If producers realise that on the introduction of a new good, consumers are purchasing more than they might be expected to in the long run, producers have an incentive to keep on introducing new goods forcing consumers to experiment. Of course, over time consumers will learn that they are not sampling from a stationary distribution. The problems that we have considered in this thesis is that the underlying distribution of the random variable remains stationary. This is not to say that consumer behaviour is not affected when the environment is adaptive and non-stationary, only that this thesis has not considered this problem.

In Chapter 7, we extended the model developed in Chapter 4 to a many period environment. We were concerned with examining the movement over time of the optimum demand for a new good, given that the consumer had been learning about it, and further would be able to learn more about it in the future. The major conclusion of that chapter was that, for the assumed observations that the consumer made, the profile of purchases increased over time such that when aggregated over a number of consumers, it produced a sigmoid diffusion curve. The shape of this curve fits with the expirical evidence on the diffusion of new products. Perhaps, more importantly, the chapter also provided a testable hypothesis on the shape of the demand functions along the life cycle. The hypothesis is that as the consumer's desire to sample the good declines over the life cycle, the elasticity of demand for the new good increases. This is because early on in the life of a product, consumers wish to learn about its qualities and are less concerned with the price. Over time, having learnt about the products characteristics the consumer becomes more price conscious.

Finally in Chapter 8, we contrasted three different learning processes in terms of the speed of acquiring complete information. The three processes represented optimal, sub-optimal and reasonable search, with the computational costs of each process declining from the first to the last. We compared the distribution of utility losses for each process in terms of a stochastic dominance criteria, since the utility loss is a random variable depending upon the observations actually taken. Ex poste we were able to say which process was best, and we found that this depended upon the range of parameter values. There was no one global optimal process.

The work examined in this thesis could be extended in two directions. We found that over time the consumer puts greater weight on his experience as the degree of confidence in his posterior beliefs increase. However, we assumed that the initial degree of confidence was given and remained fixed. Suppose instead that ϕ changes over time, perhaps according to an adaptive process given by the idfference between μ_0 and $\bar{\varepsilon}$. Then we can envisage the situation in which the overall posterior degree of confidence falls; implying the paradoxical result that false advertising causes the consumer to sample, and hence purchase, more of the good.

This would be one extension. A second improvement would be to incorporate the effect of a series of advertising statements into the model, of both the firm itself and also a rival. This would also allow us to consider the role of the rate of forgetfulness of advertising messages. How these ideas could be implemented was suggested in Chapter 3, and this could prove a fruitful avenue for further research.

Bibliography

Aitchison, J. and J.A.C. Brown (1957), The Lognormal Distribution, Cambridge University Press.

Akerlof, G.A. (1970), "'The Market for Lemons' : Quality Uncertainty and the Market Mechanism", Quarterly Journal of Economics, Vol.84, No.3 (August).

Arrow, K. (1962), Economic Implications of Learning by Doing, Review of Economic Studies, Vol.29.

Bain, J.S. (1956), Barriers to New Competition. Cambridge, Mass. Harvard University Press.

Baker, M.J. (1975), Marketing New Industrial Products, Macmillan Press.

Basman, R.L. (1956), "A Theory of Demand with Variable Consumer Preferences", Econometrica, Vol.24.

Borch, K., (1969), "A Note on Uncertainty and Indifference Curves", Review of Economic Studies, Vol.36.

Braithwaite, D. (1928), "The Economic Effects of Advertising", Economic Journal, Vol.38.

Chiplin, B. and B. Sturgess,(1981), The Economics of Advertising. Holt, Rinehart and Winston.

Cyert, R.M., M.H. DeGroot and C.A. Holt, (1978), Sequential Investment Decisions with Bayesian Learning, Management Science.

Davies, S. (1979), The Diffusion of Process Innovations, Cambridge University Press.

Deaton, A. and J. Muelbauer, (1980), Economics and Consumer Behaviour, Cambridge University Press.

DeGroot, M. (1970), Optimal Statistical Decisions.

Dixit, A.K. and V. Norman (1978), "Advertising and Welfare", The Bell Journal of Economics, Vol.9, No.1, (Spring).

Dixit, A.K. and V. Norman (1979), "Advertising and Welfare : reply", The Bell Journal of Economics, Vol.10.

Dreze, J.H. and F. Modigliani, (1972), Consumption Decision Under Uncertainty', Journal of Economic Theory, Vol.5, No.3.

Economists Advisory Group (1967), The Economics of Advertising, London: Advertising Association.

Feldstein, M.S. (1969), "Mean-variance Analysis in the Theory of Liquidity Preference and Portfolio Selection", Review of Economic Studies, Vol.36.

Ferguson, J. (1974), Advertising and Competition: Theory, Measurement, Fact, Cambridge: Ballinger Publishing Co.

Fisher, F.M. and J.J. McGowan, (1979), "Advertising and Welfare: Comment". The Bell Journal of Economics, Vol.10.

Fulop, C. (1981), Advertising, Competition and Consumer Behaviour, Holt, Rinehart and Winston.

Galbraith, J.K. (1967), The New Industrial State, Boston.

Gastwirth, J.L. (1976), "On Probabilistic Models of Consumer Search for Information", Quarterly Journal of Economics, Vol.90, No.1.

Gort, M. and S. Klepper, (1982), "Time Paths in the Diffusion of Product Innovations', Economic Journal, Vol.92.

Gould, J. (1974), "Risk, Stochastic Preference and the Value of Information", Journal of Economic Theory, Vol.8, No.1.

Grossman, S.J., R.E. Kihlstrom and L.J. Mirman, (1977), "A Bayesian Approach to the Production of Information and Learning by Doing", Review of Economic Studies, Vol.44.

Hadar, J. and W.R. Russel, (1969), "Rules for Ordering Uncertain Prospects", American Economic Review, Vol.59, No.1.

Hanoch, G. and C. Levy (1969), "Efficiency Analysis of Choices Involving Risk", Review of Economic Studies, Vol.36.

Hey, J.D. (1981), "Are Optimal Search Rules Reasonable? And Vice Versa? (And Does It Matter Anyway)", Journal of Economic Behaviour and Organisation, Vol.2.

Hicks, J.R. (1962), Economic Theory and the Evaluation of Consumer's Wants, Journal of Business, Vol.35.

Joyce, T. (1963),"Brand Images".

Joyce, T. (1967), "Advertising", both in A.S.C. Ehrenberg and F.G. Pyatt, (1971), Consumer Behaviour. Penguin.

Kihlstrom, R.E. (1974a), "A Bayesian Model of Demand for Information about Product Quality", International Economic Review, Vol.15, No.1.

Kihlstrom, R.E. (1974b), A General Theory of Demand for Information, Journal of Economic Theory, Vol.8, No.4.

Klein, B. and K.B. Leffler (1981), "The Role of Market Forces in Assuring Contractual Performance", Journal of Political Economy, Vol.89, No.4.

Kohn, M.G. and S. Shavell (1974), "The Theory of Search", Journal of Economic Theory, Vol.9.

Kotowitz, Y. and F. Mathewson (1979), "Advertising, Consumer Information and Product Quality", The Bell Journal of Economics, Vol.10.

Lancaster, K.J. (1966), "A New Approach to Consumer Theory", Journal of Political Economy, Vol.74, No.2.

Lancaster, K. (1975), Socially Optimal Product Differentiation, American Economic Review, Vol.65.

Manning, R., and P.B. Morgan, (1982), "Search and Consumer Theory", Review of Economic Studies, Vol.49.

Marschak, J. (1954), "Towards an Economic Theory of Organisation" in R.M. Thrall, C.H. Coombs and R.L. Davis, Eds., Decision Processes, Wiley, New York.

Marshall, A. (1920), Industry and Trade (3rd edn.) Macmillan.

Marshall, A. (1920), Principles of Economics (8th edn.) Macmillan.

McCall, J.J. (1970), "The Economics of Information and Job Search" Quarterly Journal of Economics, Vol.84.

Midgley, D.F. (1977), Innovations and New Product Marketing, Croom Helm, London.

Mishan, E. (1969), 21 Popular Economic Fallacies. Penguin.

Mood, A.M., F.A. Graybill and D.C. Boes (1974), "Introduction to the Theory of Statistics", McGraw-Hill, 3rd edn.

Muth, R.F. (1966), "Household Production and Consumer Demand Functions", Econometrica, Vol.34.

Nelson, P. (1970), "Information and Consumer Behaviour", Journal of Political Economy, Vol.78, No.2.

Nelson, P. (1974), "Advertising as Information", Journal of Political Economy, Vol.82, No.4.

Pigou, A.C. (1932), The Economics of Welfare (4th edn) Macmillan.

Rothschild, M. (1974a), "Searching for the Lowest Price when the Distribution of Prices is Unknown", Journal of Political Economy, Vol.82.

Rothschild, M. (1974b), "A Two-Armed Bandit Theory of Market Prices", Journal of Economic Theory, Vol.9.

Samuelson, R.A. (1970), "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments", Review of Economic Studies, Vol.37.

Sandmo, A. (1970), "The Effect of Uncertainty on Savings Decisions", Review of Economic Studies, Vol.37.

Scott, R.C. and P.A. Horvath (1980), "On the Direction of Preference for Moments of Higher Order than the Variance", Journal of Finance, Vol.35, No.4.

Stigler, G.J. (1961), "The Economics of Information", Journal of Political Economy, Vol.69.

Stoneman, P. (1980), "The Rate of Imitation, Learning and Profitability", Economic Letters, Vol. 6.

Stoneman, P. (1981), "Intra-firm Diffusion, Bayesian Learning and Profitability", Economic Journal, Vol.91.

Telser, L.G. (1966), "The Supply and Demand for Advertising Messages", American Economic Review, No.56.

Theil, H. (1978), Introduction to Econometrics, Englewood Cliffs, N.J.: Prentice-Hall.

Tintner, G. (1952), "Complementarity and Shifts in Demand", *Metroeconomica*, Vol.4.

Tobin, J. (1958), "Liquidity Preference as a Behaviour Towards Risk", *Review of Economic Studies*.

Tonks, I. (1983), "Bayesian Learning and the Optimal Investment Decision of the Firm", *Economic Journal*, A.U.T.E. Conference Edition.

Tonks, I. (1983), "A Bayesian Approach to the Production of Information with a Linear Utility Function", *Review of Economic Studies*, forthcoming.